



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number 135068

TO: Patricia Duffy
Art Unit: 1645
Location: REM/3B05/3C18
Serial Number: 09/674277

Tuesday, June 07, 2005

From: Beverly Shears
Location: Biotech-Chem Library
REM 1A54
Phone: 571-272-2528
beverly.shears@uspto.gov

Search Notes

Patricia,

Seq. IDs 2 and 21-27 are nucleic acids. The nucleic acid databases were searched. If you would like to have the amino acid databases searched, pls. contact me.

Beverly

This Page Blank (uspto)

From: Chan, Christina
Sent: Tuesday, March 08, 2005 5:02 PM
To: Duffy, Patricia; STIC-Biotech/ChemLib
Subject: RE: Sequence search RUSH, amendment due this biweek.

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Duffy, Patricia
Sent: Tuesday, March 08, 2005 2:21 PM
To: Chan, Christina
Subject: Sequence search RUSH, amendment due this biweek.
Importance: High

Chris,

please approve the following rush search. Amendment due this biweek.

In re: 09/674,277

Please search residues 400-407 of SEQ ID NO:1, SEQ ID NO:1, SEQ ID NOS:10, 11, 12, 13, 18, 19, 20.

This is a PCT case with unity of invention.

These are all short nucleic acid sequences.

Please perform an interference and commercial database hit.

Please print out top 25 hits in each category.

Thanks,

Patricia A. Duffy, Ph.D.

Art Unit 1645, Remsen 3B05

571-272-0855

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 08:55:10 ; Search time 5404 Seconds
(without alignments)
10589.491 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgcaggagatggaataaa.....ttttactttttttctctgcag 1181

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hrg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1181	100.0	1181	6	BD205212 Nucleotid
2	1181	100.0	1181	6	AX011298 Sequence
3	1166.2	98.7	7395	1	EC74P0157 Y11275 E.coli 7.4
4	1166.2	98.7	7395	6	AX191728 Sequence
5	1166.2	98.7	92077	1	AF074613 Escherich
6	1166.2	98.7	92077	6	AX191727 Sequence
7	1166.2	98.7	92721	1	AB011549 Escherich
8	1166.2	98.7	92721	6	AX191725 Sequence
9	1163	98.5	5612	1	AF043470 Escherich
10	328.6	27.8	213494	1	AF386526 Shigella
11	328.6	27.8	221618	1	AF386526 Shigella
12	328.6	27.8	221851	1	AF348706 Shigella
13	327	27.7	3815	1	AF206446 Shigella
14	246.6	20.9	3500	1	AF134403 Escherich
15	159.4	13.5	1642	1	D11025 Shigella fl
16	159.4	13.5	1642	6	E06635 Shigella vi
17	154.8	13.1	66714	1	AF326777 Shigella
18	93	7.9	300774	1	AE016912 Chromobac
19	64.2	5.4	6540	1	L38424 Bacillus su

Duffy P.
091674277
Seg. 10 2 21-27

20	64.2	5.4	10801	6	AX100832 Sequence
c 21	64.2	5.4	23775	1	BACPIA
22	64.2	5.4	201375	1	BSUR0012
c 23	63	5.3	348251	1	EX640423 Bordetell
c 24	61.4	5.2	347356	1	EX640437 Bordetell
25	60.2	5.1	300885	1	AE012559 Xylella f
c 26	59.2	5.0	13334	1	AE003927 Xylella f
27	57.2	4.8	110000	1	Continuation (25 o
28	57.2	4.8	306050	1	AX248341 Mycobacte
29	57.2	4.8	348247	1	EX842579 Mycobacte
c 30	57	4.8	14734	1	AF498417 Pseudomon
c 31	57	4.8	16736	1	AF035937 Pseudomon
c 32	56.8	4.8	202301	1	AF035937 Pseudomon
c 33	55.4	4.7	14738	1	AE017286 Desulfovi
c 34	54.8	4.6	237221	1	AF540991 Pseudomon
c 35	53.4	4.5	110000	1	AP003599 Nostoc sp
c 36	52.2	4.4	10197	1	AE017180_24
37	52.2	4.4	12006	1	AE009285 Agrobacte
c 38	52.2	4.4	110000	1	Continuation (13 o
c 39	50.6	4.3	20693	1	SMU52844
c 40	50.6	4.3	294300	1	AE017269 Bacillus
c 41	50.2	4.3	110000	1	Continuation (15 o
c 42	50.2	4.3	110000	1	Continuation (15 o
c 43	50.2	4.3	110000	1	Continuation (15 o
c 44	50.2	4.3	110000	1	Continuation (15 o
c 45	50.2	4.3	110000	1	Continuation (16 o

ALIGNMENTS

RESULT 1	BD205212	1181 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).				
DEFINITION	BD205212				
ACCESSION	BD205212				
VERSION	BD205212.1 GI:33014982				
KEYWORDS	JP 2002512813-A/2.				
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 1181)				
AUTHORS	Frechon,D.T.M., Laure,F.C. and Thierry,D.				
TITLE	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)				
JOURNAL	Patent: JP 2002512813-A 2 08-MAY-2002;				
COMMENT	BIORAD PASTEUR OS Unidentified PN JP 2002512813-A/2 PD 08-MAY-2002 PF 27-APR-1999 JP 2000546051 PR 28-APR-1998 FR 98/05329 PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC Strandedness: Double; CC Topology: Linear; CC Nucleotide sequence for detecting enterohemorrhagic CC Escherichia coli CC (EHEC). FH Key Location/Qualifiers FT source 1..1181 /organism='Unidentified'. FT Location/Qualifiers 1..1181 /organism='unidentified' /mol_type='genomic DNA' /db_xref='taxon:32644'				
FEATURES	source				
ORIGIN	Query Match 100.0%; Score 1181; DB 6; Length 1181; Best Local Similarity 100.0%; Pred. No. 0;				

Db	481	TTTCCGAAGGATGACACACACTCGGAATCCGGTAGACGGGTAAACGCTGGGAAAGGG	540
Qy	541	CACCTGCCATTAAACACATCTCCGCTCATCTCCAGGTGTTCTGCTGCTGACGACAGCGTG	600
Db	541	CACCTGCCATTAAACACATCTCCGCTCATCTCCAGGTGTTCTGCTGCTGACGACAGCGTG	600
Qy	601	CTTCGTATTCTTACGCGCGCGCGCCACACGAGGACGAGGAAATGATTTCCCTTCATCT	660
Db	601	CTTCGTATTCTTACGCGCGCGCGCCACACGAGGACGAGGAAATGATTTCCCTTCATCT	660
Qy	661	TCAGCTGATACAAATACACGAGCAGATAAATTCATGCTCTTTTTCGGGACGTAGCATCCCCA	720
Db	661	TCAGCTGATACAAATACACGAGCAGATAAATTCATGCTCTTTTTCGGGACGTAGCATCCCCA	720
Qy	721	CCTGAACGATAAGCGGAACATTGCTCTGATGACAGCCCGCGGTGATGACAGGGGTA	780
Db	721	CCTGAACGATAAGCGGAACATTGCTCTGATGACAGCCCGCGGTGATGACAGGGGTA	780
Qy	781	ACGGTCGATGGCTTCATTTATGCAATGCGGGCCAGTCGAAACCCGGTGGAAATACCGTTA	840
Db	781	ACGGTCGATGGCTTCATTTATGCAATGCGGGCCAGTCGAAACCCGGTGGAAATACCGTTA	840
Qy	841	CCGGTGCTCTGACACCTTCGCGCATCAGATGCGGCATCATGGGTGAGATAGGCACACAA	900
Db	841	CCGGTGCTCTGACACCTTCGCGCATCAGATGCGGCATCATGGGTGAGATAGGCACACAA	900
Qy	901	TGAATCACACAGATAATTACAGGAAACGTTCTGGTCTTACGGGTGATGAGTGTGTTTT	960
Db	901	TGAATCACACAGATAATTACAGGAAACGTTCTGGTCTTACGGGTGATGAGTGTGTTTT	960
Qy	961	GTCGACAAATAGTGAACGGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTACTGT	1020
Db	961	GTCGACAAATAGTGAACGGGTGACAGCATATCAGACGGCTCAGTCTGCTATATTACTGT	1020
Qy	1021	CATGCCCATATGACAGATGACAGTTCAGTTCATTCATTCATTCATTCATTCATTCAT	1080
Db	1021	CATGCCCATATGACAGATGACAGTTCAGTTCATTCATTCATTCATTCATTCATTCAT	1080
Qy	1081	TGAGGATGGAAGGAAGTGAAGGCTGTTCTGAAAGGAATAAAGTGACATCATGCCCTC	1140
Db	1081	TGAGGATGGAAGGAAGTGAAGGCTGTTCTGAAAGGAATAAAGTGACATCATGCCCTC	1140
Qy	1141	TTTTCTGGCTTCGGAGCAATTTTACTTTTTTCTGCGAG	1181
Db	1141	TTTTCTGGCTTCGGAGCAATTTTACTTTTTTCTGCGAG	1181
RESULT 3			
EC74P0157/c			
LOCUS	EC74P0157	7395 bp	DNA linear BCT 07-JAN-1999
DEFINITION	E.coli 7.4 kb DNA from plasmid p0157.		
ACCESSION	Y11275		
VERSION	Y11275.1	GI:4127812	
KEYWORDS	mabB gene; unidentified protein.		
SOURCE	Escherichia coli		
ORGANISM	Escherichia coli		
	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.		
REFERENCE	1	Brunder, W.	
AUTHORS	Unpublished		
JOURNAL	2 (bases 1 to 7395)		
REFERENCE	2	Brunder, W.	
AUTHORS	Direct Submission		
TITLE	Submitted (14-FEB-1997) W. Brunder, Universitaet Wuerzburg,		
JOURNAL	Institut fuer Hygiene und Mikrobiologie, Josef-Schneider-Strasse 2,		
	D- 97080 Wuerzburg, FRG		
COMMENT	Related sequences X97542, M77039, X02311.		
FEATURES	Location/Qualifiers		
source	1..7395	/organism="Escherichia coli"	
		/mol_type="genomic DNA"	
		/strain="EDL933"	

/aerotype="O157:H7"	
/db_xref="taxon:562"	
/clone="psm7"	
/plasmid="p0157"	
774..1595	
/codon_start=1	
/transl_table=11	
/product="hypothetical protein"	
/protein_id="CAA72138.1"	
/db_xref="GI:4127813"	
/db_xref="GOA:O85100"	
/db_xref="UniProt/TREMBL:O85100"	
/translation="MLNARHLPVLMYHVSRCPLVLTLSPETCEOMKWLADNNRTV	
TABLSYFQGGTLPRKSVMLTDDGLDNWLRWPKVKYNLHAHIFLITIGKGN	
VSRQEQEYSHRDCERLIADNSDEVMRWSEVRMDSLGVFHVHHTHGHKWRDL	
VRAEQCRIMKEDIIVKQCLTEKLGCSHLCPGEGYNNRDIINLAGKLGFSLYTT	
ERRMCPENGSLRIGRISTKERHSGWLKRLFFYTTPLFFSVLALHKGRLPDN"	
1595..2701	
/codon_start=1	
/transl_table=11	
/product="hypothetical protein"	
/protein_id="CAA72139.1"	
/db_xref="GI:4127814"	
/db_xref="GOA:O82935"	
/db_xref="UniProt/TREMBL:O82935"	
/translation="MMKILFTSSSDIGGOELQALQMTALQKQHSVLLACREKSI	
APEARKHGHDVTFIPFRSLHLPISILRLRIIGEPKPDLVICHSHDSNLAGLSRL	
CHFSIVROKTYITRKTFTSLNYLCDFIVPSSAMMAHLMAGVTRPTVTIIPGDFM	
PLHNAMRPLPHIHAWAASADNVPLIVQVGNLPEKGEHFLVLYQKMEGKSPR	
WLVVAGREYEREARLPQOTHELGMSGDVLMAGALFPALPVYRIASVVVMSSEAEFCM	
VLABASGVPIVASETGIPDVIQKNTGTLLPVGVDVSANTGALRDLFLSRPFRFM	
ASASRDIYRFIDINATQIIIVSLSAQAKGKCNR"	
2791..4512	
/codon_start=1	
/transl_table=11	
/product="hypothetical protein"	
/protein_id="CAA72140.1"	
/db_xref="GI:4127815"	
/db_xref="GOA:O9R768"	
/db_xref="UniProt/TREMBL:O9R768"	
/translation="MHLNTGQNRPTFSWAGLWAIYVFGPFSTLLOVIFSSGYSGTN	
GIRDSLLFSCLMLIPVLPDRKIIAAVVGFILWGTSALACVLYFGHERPSQSVLF	
VMPETNAREAGYSFQYFSLKLLLSLVYTAVSFVLRRLRPVYILPWRRIIVSFLLL	
YALLHPVVLKSLIQBELNDTLGKVASRMEAPQWQVSSYYQHQQLNALTTFLME	
NSALPLGNLRDESGRPTLVLIGESTQRMRSYGLYRRETTPELDALRKTDPGIT	
VFNVVASRPTTIALQALTFANEKNPDLYTPQSLMMNMKGQKTYITNQQTIT	
ARNTMLTVFSROTDROYVMQORTQSAEYDYNLKPFRVLNDPAKPKLIIVHLLGT	
HIKYKRYPEGGGRPDGTGHIPTGLNAKELEVYNDYDNANFVNDHVASLIKDFRAT	
APDGLLYSDHGEVYDTPPYKTQGRNEDNPTPMYTPVFLLTSEKWHAAHPRDFS	
QVDRKYSALHLHTWSDLAGLTYDGYDPTSRSLVNPQFRETTRWIGNPYKKGGLTDFD	
TLPGGSP"	
4586..5584	
/gene="mabB"	
4586..5584	
/gene="mabB"	
/codon_start=1	
/transl_table=11	
/product="MabB protein"	
/protein_id="CAA72141.1"	
/db_xref="GI:4127816"	
/db_xref="GOA:O82937"	
/db_xref="UniProt/TREMBL:O82937"	
/translation="MVPPSAVLCYHNEISRQIPVNMKNINTEFIPRNLTLCPFRYWM	
TWTGIGIICVFAMPVLPALRDLPLGLKGLMVLGRKLGARQALINLSLCFFPEYDKEKE	
NIVDAMFATAMAVLMLALSGDPKISHRIWNGLEIVKMAONNEKIVFLVPHAW	
GVDIPMLAAASGRKMAAMFHQNRNVVDYVWNSVRRRFGGKLHARDGASVRSVR	
QGWGYLLPDQHGPFSEADFATYKATLFPVIGRLSRISGARILPLFPYVQGTTHH	
LTHVSPPLAIRIQKSDAHARQINVEVNFVRPHPSQYTWILKLLKTRKEGSDPY"	
5645..5820	
/note="IS3-like sequence"	
6465..7395	
/codon_start=1	
/transl_table=11	

misc_feature

CDS

/product="hypothetical protein"
/protein_id="CAA72142.1"
/db_xref="GI:4127817"
/db_xref="GO:Q9ZAL1"
/db_xref="UniProt:TFEMBL:Q9ZAL1"
/translations="MNTKMRWRTPMKLYLSCTTILAPLAIQVFSATAADNNNSAIYF
NTSQINDLQSLAAEYKFAQSGLPAHPKREGDSQPHLSLRKSLLLVLRPVKADKTP
VOYARDNNKILGLTLYLPPSSLPDTIYHLGVPEGIDFTPHNGTKKIINTVAEVN
KLSDAGSSSIHSLTNALVEIHTANGRWVDIYLPQGPDECKMVRFFVSSAGYSTV
FYGDKVTLSVGNLTLPKYNGQWFRSGLENNRIITYAQHINSABELPAHWIVPGLNLV
IKQNLUGRLNDIKIGAPGELLITIDIGMLTTP"

ORIGIN

Query Match 98.7%; Score 1166.2; DB 1; Length 7395;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
DB 2875 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 2816

QY 71 TGTTCGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 TGTTCGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756

QY 131 CTCGGTTACCGGAAAAACCGCTGAAACAAAATTCGGGCTGAAAGAGGATCCCGCTTATCT 190
DB 2755 CTCGGTTACCGGAAAAACCGCTGAAACAAAATTCGGGCTGAAAGAGGATCCCGCTTATCT 2696

QY 191 GTTGCAATTCCTTACCTGAGTACGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 250
DB 2695 GTTGCAATTCCTTACCTGAGTACGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 2636

QY 251 ATCAAAACCGGTACTCAATATCTTCTGCGGCTGGGCTGCGCATCATCGGAGCGTTCGG 310
DB 2635 ATCAAAACCGGTACTCAATATCTTCTGCGGCTGGGCTGCGCATCATCGGAGCGTTCGG 2576

QY 311 TCGGGATAAAAAATTCGGGATGCGCGTCCATGACAGACACATCCCGCCACGGGTAAACAG 370
DB 2575 TCGGGATAAAAAATTCGGGATGCGCGTCCATGACAGACACATCCCGCCACGGGTAAACAG 2516

QY 371 CGTCCCTGTGACATCTTCTGAAATGACATCAGGATCCCGCGTGTCTCACTGGCGATAAC 430
DB 2515 CGTCCCTGTGACATCTTCTGAAATGACATCAGGATCCCGCGTGTCTCACTGGCGATAAC 2456

QY 431 GGGCAGCCCGGAGACTGACCGCTTCAGCAGTACCATACCAACAGCTTCATTTCCGAAGG 490
DB 2455 GGGCAGCCCGGAGACTGACCGCTTCAGCAGTACCATACCAACAGCTTCATTTCCGAAGG 2396

QY 491 CATGACCCACCACTGGCAATCCGGTAGACCGGTAAACCGTGGGAAAAAGGCACTGCGCAT 550
DB 2395 CATGACCCACCACTGGCAATCCGGTAGACCGGTAAACCGTGGGAAAAAGGCACTGCGCAT 2336

QY 551 TAACACATCTCCGTCATCCAGGTGTTCTGTCTGTGACGAGAGTGTCTGATTC 610
DB 2335 TAACACATCTCCGTCATCCAGGTGTTCTGTCTGTGACGAGAGTGTCTGATTC 2276

QY 611 TTCACGCCCGCGCCACACAGGACGAGCAATGATTTCCCTTCCATCTTCAGCTGATA 670
DB 2275 TTCACGCCCGCGCCACACAGGACGAGCAATGATTTCCCTTCCATCTTCAGCTGATA 2216

QY 671 CAATACACGAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAAACGAT 730
DB 2215 CAATACACGAGCATAAATTCATGTCCTTTTTCGGGACGTAGCATCCCCACCTGAAACGAT 2156

QY 731 AAGGGAAATTTGTCTGATGACGCCAGCGGTGATATGACGGGTAAACCGGTGCGCAT 790
DB 2155 AAGGGAAATTTGTCTGATGACGCCAGCGGTGATATGACGGGTAAACCGGTGCGCAT 2096

QY 791 GGCCTTCAATATGCAATGCGGCGCAGTCGAAACCGGTGGAATAACCGTTACCGGTGTCCT 850
DB 2095 GGCCTTCAATATGCAATGCGGCGCAGTCGAAACCGGTGGAATAACCGTTACCGGTGTCCT 2036

QY 851 GACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGATAGGCAACAATGAATCACA 910
DB 2035 GACACCTTCGCCCATCAGATGCGCCATCATGGGTGAGATAGGCAACAATGAATCACA 1976

QY 911 CAGATAATTCAGGAAAAAGCTTCGTCTTACCGGTGATGTAGGTTTTTGTCTGACAAT 970
DB 1975 CAGATAATTCAGGAAAAAGCTTCGTCTTACCGGTGATGTAGGTTTTTGTCTGACAAT 1916

QY 971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCGTCTGTATATTAATCTGTCATGGCCACT 1030
DB 1915 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCGTCTGTATATTAATCTGTCATGGCCACT 1856

QY 1031 ATGGCAGATGACAGCATCAGGTTTTAAATTCGCCGATAATCCCGTGAAGTCTGAGGATGGA 1090
DB 1855 ATGGCAGATGACAGCATCAGGTTTTAAATTCGCCGATAATCCCGTGAAGTCTGAGGATGGA 1796

QY 1091 AGGAGGTGAAGCTTCTCTGAAAGCAATAAAGTGACATCATGCGCTCTTTTCTGGC 1150
DB 1795 AGGAGGTGAAGCTTCTCTGAAAGCAATAAAGTGACATCATGCGCTCTTTTCTGGC 1736

QY 1151 TTCGGGAGCAATTTTACTTTTTTCTCTGCAG 1181
DB 1735 TTCGGGAGCAATTTTACTTTTTTCTCTGCAG 1705

RESULT 4
AX191728/c
LOCUS AX191728 7395 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 10 from Patent WO0149775.
ACCESSION AX191728
VERSION AX191728.1 GI:15209897
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE 1
AUTHORS Iversen, P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 10 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
Location/Qualifiers
source 1..7395
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN

Query Match 98.7%; Score 1166.2; DB 6; Length 7395;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
DB 2875 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 2816

QY 71 TGTTCGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
DB 2815 TGTTCGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756

QY 131 CTCGGTTACCGGAAAAACCGCTGAAACAAAATTCGGGCTGAAAGAGGATCCCGCTTATCT 190
DB 2755 CTCGGTTACCGGAAAAACCGCTGAAACAAAATTCGGGCTGAAAGAGGATCCCGCTTATCT 2696

QY 191 GTTGCAATTCCTTACCTGAGTACGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 250
DB 2695 GTTGCAATTCCTTACCTGAGTACGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 2636

QY 251 ATCAAAACCGGTACTCAATATCTTCTGCGGCTGGGCTGCGCATCATCGGAGCGTTCGG 310
DB 2635 ATCAAAACCGGTACTCAATATCTTCTGCGGCTGGGCTGCGCATCATCGGAGCGTTCGG 2576

QY 311 TCGGGATAAAAAATTCGGGATGCGCGTCCATGACAGACACATCCCGCCACGGGTAAACAG 370

Db 2575 TCGGGATAAAAAATCGCGCATGCGCGGTCCATGCGAGACATATCCCCACGGGTAAACAG 2516
Qy 371 CGTCCCTGTACATTTCTTGAATGACATCAGGATCCCGCCCTCTCACTCGGCGATAAC 430
Db 2515 CGTCCCTGTACATTTCTTGAATGACATCAGGATCCCGCCCTCTCACTCGGCGATAAC 2456
Qy 431 GGGCAGCGCGGAGACTGACCGTTCAGCGAGTACATACCAACCGTTCATTTCCGGAAG 490
Db 2455 GGGCAGCGCGGAGACTGACCGTTCAGCGAGTACATACCAACCGTTCATTTCCGGAAG 2396
Qy 491 CATCACACACACTGCGCAATCCGGTAGACCGGTAAACGCTGGGAAAAGGCGACCTGCGAT 550
Db 2395 CATCACACACACTGCGCAATCCGGTAGACCGGTAAACGCTGGGAAAAGGCGACCTGCGAT 2336
Qy 551 TAACACATCTCCGTCATTTCCAGGTGTCTGTCTGCTGACGACAGCGTGTCTCGTATTC 610
Db 2335 TAACACATCTCCGTCATTTCCAGGTGTCTGTCTGCTGACGACAGCGTGTCTCGTATTC 2276
Qy 611 TTACGCGCGCGCGCCAGCAGGACCGCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 2275 TTACGCGCGCGCGCCAGCAGGACCGCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 2216
Qy 671 CAATACACGAGCATAAATTCATCTCTCTTTTCGGGACGTAGCATATCCCACTGGAACGAT 730
Db 2215 CAATACACGAGCATAAATTCATCTCTCTTTTCGGGACGTAGCATATCCCACTGGAACGAT 2156
Qy 731 AAGCGGAACATTTGCTGATGATGAGCCGCGGTGGATATGCAAGGGGTAAACGCTGCGAT 790
Db 2155 AAGCGGAACATTTGCTGATGATGAGCCGCGGTGGATATGCAAGGGGTAAACGCTGCGAT 2096
Qy 791 GGCTTCATTTGCAATCGGCGCAGTCGAAACCGCGTGGATATACCGTTACCGGTGTCCT 850
Db 2095 GGCTTCATTTGCAATCGGCGCAGTCGAAACCGCGTGGATATACCGTTACCGGTGTCCT 2036
Qy 851 GACACCTTCGCGCATCAGATCGCGCATCATGSGGTGAGATAGGACACAACATGAATCACA 910
Db 2035 GACACCTTCGCGCATCAGATCGCGCATCATGSGGTGAGATAGGACACAACATGAATCACA 1976
Qy 911 CAGATAATTCAGGGAACGTTCTGTGTTCTACCGGTGATGATGATTTTGTCTGACAAT 970
Db 1975 CAGATAATTCAGGGAACGTTCTGTGTTCTACCGGTGATGATGATTTTGTCTGACAAT 1916
Qy 971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATATCTGATGCGCACT 1030
Db 1915 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATATCTGATGCGCACT 1856
Qy 1031 ATGCGAGATGACAGATCAGGTTTAAATTCGCCGATATCCGTCGAAGTCTGAGGATGGA 1090
Db 1855 ATGCGAGATGACAGATCAGGTTTAAATTCGCCGATATCCGTCGAAGTCTGAGGATGGA 1796
Qy 1091 AGGAAGGTGAAGGCTGTTCTGAAAGGAATAAAAGTGACATCATGCGCTCTTTTCTGGC 1150
Db 1795 AGGAAGGTGAAGGCTGTTCTGAAAGGAATAAAAGTGACATCATGCGCTCTTTTCTGGC 1736
Qy 1151 TTCCGAGCAATTTTACTTTTCTCTGCGAG 1181
Db 1735 TTCCGAGCAATTTTACTTTTCTCTGCGAG 1705

RESULT 5
AF074613/c
LOCUS AF074613 92077 bp DNA circular BCT 04-NOV-1998
DEFINITION Escherichia coli O157:H7 plasmid pO157, complete sequence.
ACCESSION AF074613
VERSION AF074613.1 GI:3822114
KEYWORDS Escherichia coli O157:H7
SOURCE Escherichia coli O157:H7
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 92077)
AUTHORS Burland, V., Shao, Y., Perna, N.T., Plunkett, G., Sofia, H.J. and

Blattner, F.R.
The complete DNA sequence and analysis of the large virulence
plasmid of Escherichia coli O157:H7
Nucleic Acids Res. 26 (18), 4196-4204 (1998)
98391744
PUBMED 9722640
2 (bases 1 to 92077)
Burland, V., Shao, Y., Perna, N.T., Plunkett, G. III, Sofia, H.J. and
Blattner, F.R.
Direct Submission
Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445
Henry Mall, Madison, WI 53706, USA
Location/Qualifiers
1. .92077 /organism="Escherichia coli O157:H7"
/mol_type="genomic DNA"
/strain="EDL933"
/serotype="O157:H7"
/db_xref="taxon:83334"
/lab_host="Escherichia coli C600"
/plasmid="pO157"
1. .561 /genes="finO"
/note="synonym: L7001"
1. .561 /genes="finO"
/note="97 pct identical amino acid sequence and equal
length to F101_ECOLI SW: P22707"
/codon_start=1
/transl_table=11
/product="fertility inhibition protein (conjugal transfer
repressor)"
/protein_id="AAC70069.1"
/db_xref="GI:3822115"
/translation="MAEQKRPVLTLRKTEGETPVRSRKTIINVTTPKWKVKQKLA
EKAREAEALAAKKAQALSIYVNLPTODEAVNTLKPMPGLFDGDTPLLACGIRD
VILEDVAQRNIPLSHKLRALKAITRESESYLCAMKAGACRYDEGVTEHISQEEEA
YAAARDKTRRQNRKAKELQAVLDEK"
653. .681 /notes="predicted sigma 70 promoter; score of 52%"
698. .949 /genes="L7002"
698. .949 /genes="L7002"
/note="54 pct identical (1 gap) to 23 residues of an
approx. 269 aa protein EXOK_RHIME SW: P33693"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAC70070.1"
/db_xref="GI:3822116"
/translation="MDSETVHGTVRSVTSVPEAGPLFMKSVSDAGWKQRKHGDLPLV
HPGLTGSSLPKGLNATGAAEBGCKEKSLSHYRDSG"
1025. .1052 /note="predicted sigma 70 promoter; score of 62%"
1151. .1612 /genes="L7003"
1151. .1612 /genes="L7003"
/note="98 pct identical and equal length to YFI3_ECOLI SW:
Q99342"
/codon_start=1
/transl_table=11
/product="hypothetical protein 15.6 kDa protein in finO 3'
region precursor"
/protein_id="AAC70071.1"
/db_xref="GI:3822117"
/translation="MRKIPLVLFIFSWPVLCAIHGRVVRVLDGDTIEVWDSKAVR
IRLVNTDAPEKKQDYGWSTDMKMSLVAGKTVTVTFQRDRYGRMLGQVYAPDGMVN
QPMVRAGAARVVEQYNTDPLVPLVQNEARQQKRLMSADPVPFPWIRHRK"
1390. .1419 /note="predicted sigma 70 promoter; score of 56%"
1658. .1867

```
/gene="L7004"
1658. .1867
/gene="L7004"
/notes="55 pct identical (0 gaps) to 66 residues of an
approx. 72 aa protein HHA_ECOLI SW: P23870"
/codon_start=1
/transl_table=11
/product="putative hemolysin expression modulating
protein"
/protein_id="AAC70072.1"
/db_xref="GI:3822118"
/translation="MEKTOEWLYQIARRCSSVNTLEKLIHKRDSLSSTSERESFNAA
DRLAELITGKLYDRIPKEIKWVR"
1899. .1927
/notes="predicted sigma 70 promoter; score of 56%"
1905. .2243
/gene="L7005"
1905. .2243
/gene="L7005"
/notes="98 pct identical (0 gaps) to 57 residues of an
approx. 200 aa protein; plasmid R100 miniplasmid pSM1 ORF
4, TRSW: Q52340"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC70073.1"
/db_xref="GI:3822119"
/translations="MKLIIFILIVLIIAALLIRILRSVNOHSPLLMOLHAAAGIRTG
AERILSGGEVDASHLPERRRRDILLEVLRLQIGIFLLRSHDARKLLQMTGEWLTTG
AAQSPHRS"
complement (2240. .2327)
/notes="initiation site; Escherichia coli plasmid R100 ssIB
gene; 99 pct identical to ssIB locus R10SSIB accession
D90185"
2416. .2421
/notes="100 pct identical (0 gaps) to the -35 region at
23. .28 locus ECNR1REP accession X02302"
2439. .2445
/notes="100 pct identical (0 gaps) to the -10 region at
46. .52 locus ECNR1REP accession X02302"
2473. .2478
/notes="100 pct identical (0 gaps) to RBS at 80. .85 locus
ECNR1REP accession X02302"
2483. .2737
/gene="cpb2"
/notes="synonym: L7006"
2483. .2737
/gene="cpb2"
/notes="98 pct identical amino acid sequence and equal
length to CPB2_ECOLI SW: P03847"
/codon_start=1
/transl_table=11
/product="CpB protein (RepA2 protein)"
/protein_id="AAC70074.1"
/db_xref="GI:3822120"
/translations="MSQTENAVTSSSGAKAYRKGNPLSDAERQRLSVARKASFKEV
KVLEPKYKAMLQMCHEDGLTQAEVLITAIKSEAQRVCV"
2754. .2782
/notes="predicted sigma 70 promoter; score of 69%"
2973. .3047
/gene="L7007"
2973. .3047
/gene="L7007"
/notes="91 pct identical (0 gaps) to 24 residues of an
approx. 24 aa protein uORF P30REPFIC, accession M16167,
translationally coupled to replication initiation protein;"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70075.1"
/db_xref="GI:3822121"
/translation="MLGKVQDFFLCSLILLRIVSAGWCD"
3032. .3035
```

```
gene
/notes="100 pct identical (0 gaps) to RBS at 641. .644 locus
ECNR1REP accession X02302"
3040. .3897
/gene="rep2"
/notes="synonym: L7008"
3040. .3897
/gene="rep2"
/notes="96 pct identical amino acid sequence and equal
length to REP2_ECOLI SW: P03066"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70076.1"
/db_xref="GI:3822122"
/translation="MTDLQOYYVROKPNPVPTREGAGTLKFCBKLMEKAVGFTSR
FDPATHVAHARSKGLRRMPVLRRAIDALLQGLCFHYDPLANRVQCSITTLAIBCG
LATESNAGKLSITRATRALTFILSELGITYQTEYDPLIGCIPTDLTFPTPALADLV
SADAVAAARSRVWENRURKQGLTGMDELIAKAWFVRFRFRSYQTELKSRGK
RARARRADGRERQDIVTLVKRLQTLREISGRFTANREAVKREVERVRKERMILLSRNR
YSRLATASP"
3847. .4095
/standard_name="CIS"
/notes="88 pct identical to locus ECCIS accession X12587,
required for cis-activation of oriR by the replication
initiation protein"
4072. .4080
/notes="dnaA site; 100 pct identical (0 gaps) to locus
ECNR1REP at (1682. .1690) accession X02302"
4084. .4232
/standard_name="oriR"
/notes="89 pct identical to oriR (1094. .1242); minimum
segment for replication of E. coli IncFII plasmid NR1
ECREPA1 X12776"
/direction=right
4169. .4197
/notes="predicted sigma 70 promoter; score of 56%"
4259. .4453
/gene="L7009"
4259. .4453
/gene="L7009"
/notes="95 pct identical to (0 gaps) 64 residues of a 128
aa protein REPA4 locus ECRS1 accession V00351"
/codon_start=1

Query Match 98.7%; Score 1166.2; DB 1; Length 92077;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 TGGAAAAAGCCAAATAAAAAATTTGCCATCCAGCGGCTCCAGCTGAAAGTAGGCC 70
Db 19424 TGGAAAAAGCCAAATAAAAAATTTGCCATCCAGCGGCTCCAGCTGAAAGTAGGCC 19365

Qy 71 TGTTCGTCCGGTATTAAATGCATTGACCGTCCCGTATTTAAACAATGTGATAATTA 130
Db 19364 TGTTCGTCCGGTATTAAATGCATTGACCGTCCCGTATTTAAACAATGTGATAATTA 19305

Qy 131 CTCGGTTACCGAAAAACCGCTGAACAAAAATTCGGGCTGAAAAGAGGATCCCGCTTATCT 190
Db 19304 CTCGGTTACCGAAAAACCGCTGAACAAAAATTCGGGCTGAAAAGAGGATCCCGCTTATCT 19245

Qy 191 GTTGCAATTTCCCTTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTGAAT 250
Db 19244 GTTGCAATTTCCCTTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTGAAT 19185

Qy 251 ATCAACCGGTACTCAATATCTTCTGCGGCTGGCTGCCATCATCCGGAAGGTTCCCG 310
Db 19184 ATCAACCGGTACTCAATATCTTCTGCGGCTGGCTGCCATCATCCGGAAGGTTCCCG 19125

Qy 311 TCGGATAAAAAATCCGCGAGTCGCGGCTCCATGACAGACACATCCCCACCGGTAAACAG 370
Db 19124 TCGGATAAAAAATCCGCGAGTCGCGGCTCCATGACAGACACATCCCCACCGGTAAACAG 19065

Qy 371 CGTCCCTGTACACATTTCTTGAATGATCATCAGGGATCCCGCGCTCTCACTGGCGATAAC 430
```


source		1..92077		/organism="Escherichia coli"		/mol_type="unassigned DNA"		/db_xref="taxon:562"	
ORIGIN		Query Match		98.7%; Score 1166.2; DB 6; Length 92077;					
		Best Local Similarity		99.7%; Pred. No. 0;					
		Matches 1168; Conservative		0; Mismatches 3; Indels 0; Gaps 0;					
Qy	11	TGGA	AAAAAAGCCAAAATAAAAAATTC	CCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC	70				
Db	19424	TGGA	AAAAAAGCCAAAATAAAAAATTC	CCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC	19365				
Qy	71	TGTT	CTGTCCGGTATTTAAATGCAATG	ACCGTCCCGTATTTAAACAATGTGATAAATTA	130				
Db	19364	TGTT	CTGTCCGGTATTTAAATGCAATG	ACCGTCCCGTATTTAAACAATGTGATAAATTA	19305				
Qy	131	CTCG	GTTACCGGAAAACCGCTGAACAAA	ATTCGGGCTGAAAAGAGGATCCGCCGTTATCT	190				
Db	19304	CTCG	GTTACCGGAAAACCGCTGAACAAA	ATTCGGGCTGAAAAGAGGATCCGCCGTTATCT	19245				
Qy	191	GTTC	GCATTTCCCTTACGCTGACTAGC	CAGAGACAAATGATCTGTGCCGTTCTGTGTAAT	250				
Db	19244	GTTC	GCATTTCCCTTACGCTGACTAGC	CAGAGACAAATGATCTGTGCCGTTCTGTGTAAT	19185				
Qy	251	ATCA	AAACCGGTACTCAATATCTTCT	TCTGGCGCTGGTGCATCATCCGGAAGCGTTC	CGG 310				
Db	19184	ATCA	AAACCGGTACTCAATATCTTCT	TCTGGCGCTGGTGCATCATCCGGAAGCGTTC	CGG 19125				
Qy	311	TCGG	GATAAAAAATTCGGGAGTCCG	CGGTCCATGACAGACACATCCCGCCAGGGTAACAG	370				
Db	19124	TCGG	GATAAAAAATTCGGGAGTCCG	CGGTCCATGACAGACACATCCCGCCAGGGTAACAG	19065				
Qy	371	CGTC	CCCTGTCCACATTTCTTCTGA	ATGACATCAGGGATCCCGCGCTCTCACTGGCGGATAAC	430				
Db	19064	CGTC	CCCTGTCCACATTTCTTCTGA	ATGACATCAGGGATCCCGCGCTCTCACTGGCGGATAAC	19005				
Qy	431	GGGC	CACGCGGAGACTGACGCTTC	CAGCAGTACCATAACCAACGCTTCATTTTCCGAAGG	490				
Db	19004	GGGC	CACGCGGAGACTGACGCTTC	CAGCAGTACCATAACCAACGCTTCATTTTCCGAAGG	18945				
Qy	491	CATG	ACACACACTGGCAATCCGGT	AGACCGGTAACTGGGAAAAGGGCACCTGCGCAT	550				
Db	18944	CATG	ACACACACTGGCAATCCGGT	AGACCGGTAACTGGGAAAAGGGCACCTGCGCAT	18885				
Qy	551	TAA	CACATCTCCGCTCATTTCC	AGGTGTTCTGTCTGTGACGACAGCTGCTTCGTATTC	610				
Db	18884	TAA	CACATCTCCGCTCATTTCC	AGGTGTTCTGTCTGTGACGACAGCTGCTTCGTATTC	18825				
Qy	611	TTCA	CGCCGCGGCCACCAACGAG	CCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA	670				
Db	18824	TTCA	CGCCGCGGCCACCAACGAG	CCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA	18765				
Qy	671	CAAT	ACACGAGCATAAATTCATG	CTCTTTCGGGACGTAGCATCCCACTGAAACGAT	730				
Db	18764	CAAT	ACACGAGCATAAATTCATG	CTCTTTCGGGACGTAGCATCCCACTGAAACGAT	18705				
Qy	731	AAGC	GGAACATTTGCTCTGATGC	AGCCAGCGCTGGATATGCAGGGGTAAACGGTCCGAT	790				
Db	18704	AAGC	GGAACATTTGCTCTGATGC	AGCCAGCGCTGGATATGCAGGGGTAAACGGTCCGAT	18645				
Qy	791	GGCT	TTCATTAATGCAATCGGCG	CCAGTCGAAACCCGGTGGAATTAACCGGTTACCGGTTGCTCT	850				
Db	18644	GGCT	TTCATTAATGCAATCGGCG	CCAGTCGAAACCCGGTGGAATTAACCGGTTACCGGTTGCTCT	18595				
Qy	851	GAC	ACCTTCGGCATCAGATCGC	CAATCATATGGGTAGATAGGACAACAATGAAATCACA	910				
Db	18584	GAC	ACCTTCGGCATCAGATCGC	CAATCATATGGGTAGATAGGACAACAATGAAATCACA	18525				
Qy	911	CAG	ATAATTCAGGAAAACGTTCT	GGTCTTACGGGTGATGTAGGTTTTTGTCTGACAAT	970				
Db	18524	CAG	ATAATTCAGGAAAACGTTCT	GGTCTTACGGGTGATGTAGGTTTTTGTCTGACAAT	18465				
Qy	971	AGT	GAAGCGGTGACAGCATATC	AGACGGCTCAGTCTGTATATTACTGTCTATGGCCACT	1030				
Db	18464	AGT	GAAGCGGTGACAGCATATC	AGACGGCTCAGTCTGTATATTACTGTCTATGGCCACT	18405				
Qy	1031	ATG	GAGATGACAGATCAGGTTT	AAATTTCCCGATAATCCGTGAAAGTCTGAGGATGGA	1090				
Db	18404	ATG	GAGATGACAGATCAGGTTT	AAATTTCCCGATAATCCGTGAAAGTCTGAGGATGGA	18345				
Qy	1091	AGG	AGGTGAAGGCTGTTCTCTG	AAGGAATAAAGTCACATCATGCCCTCTTTTCTGGC	1150				
Db	18344	AGG	AGGTGAAGGCTGTTCTCTG	AAGGAATAAAGTCACATCATGCCCTCTTTTCTGGC	18285				
Qy	1151	TTCC	GGAGCAATTTTACTTTTTT	CTCTGCGAG 1181					
Db	18284	TTCC	GGAGCAATTTTACTTTTTT	CTCTGCGAG 18254					
RESULT 6									
AX191727/c									
LOCUS	AX191727	92077 bp	DNA	linear	PAT 15-AUG-2001				
DEFINITION	Sequence 9 from Patent WO0149775.								
ACCESSION	AX191727								
VERSION	AX191727.1 GI:15209896								
KEYWORDS									
SOURCE									
ORGANISM	Escherichia coli								
REFERENCE	1								
	Iversen, P.L.								
	Antisense antibacterial cell division composition and method								
	Patent: WO 0149775-A 9 12-JUL-2001;								
	Avi Biopharma, Inc. (US)								
FEATURES	Location/Qualifiers								

Qy 971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTTGCTATATTAATCTGTCATGGCCACT 1030
|
Db 18464 ACTGAAGCGGTGACAGCATATCAGACGGCTCAGTCTTGCTATATTAATCTGTCATGGCCACT 18405
|
Qy 1031 ATGGCAGATACACAGATCAGGTTTAAATTTCCCGATTAATCCGTCGAAGTCTGAGGATGGA 1090
|
Db 18404 ATGGCAGATACACAGATCAGGTTTAAATTTCCCGATTAATCCGTCGAAGTCTGAGGATGGA 18345
|
Qy 1091 AGGAAGGTGAAGCGTGTCTCTGAAAGAAATAAAGTGACATCATGCCCTCTTTTCTGGC 1150
|
Db 18344 AGAAGGTGAAGCGTGTCTCTGAAAGAAATAAAGTGACATCATGCCCTCTTTTCTGGC 18285
|
Qy 1151 TTCGGAGCAATTTTACTTTTTTCTCTGCGAG 1181
|
Db 18284 TTCGGAGCAATTTTACTTTTTTCTCTGCGAG 18254
|

RESULT 7
AB011549/c
LOCUS AB011549 92721 bp DNA circular BCT 27-APR-1999
DEFINITION Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION AB011549
VERSION AB011549.2 GI:4589740
KEYWORDS ToxR-regulated lipoprotein; tagA.
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
1 (sites)
Makino, K., Ishii, K., Yasunaga, T., Hattori, M., Yokoyama, K.,
Yutsudo, H. C., Kubota, Y., Yamaichi, Y., Iida, T., Yamamoto, K.,
Honda, T., Han, C. G., Ohtsubo, E., Kasamatsu, M., Hayashi, T., Kuhara, S.
and Shinagawa, H.
Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
outbreak
DNA Res. 5 (1), 1-9 (1998)
98290540
9628576
2 (bases 1 to 92721)
Makino, K.
Direct Submission
Submitted (24-FEB-1998) Kozo Makino, Research Institute for
Microbial Diseases, Osaka University, Molecular Microbiology;
Yamadaoka, 3-1, Suita, Osaka 562, Japan
[E-mail: makino@bks01.biken.osaka-u.ac.jp, Tel: 81-6-879-8318,
Fax: 81-6-879-8320]
On Apr 20, 1999 this sequence version replaced gi:3336997.

FEATURES
source
1. .92721
Location/Qualifiers
/organism="Escherichia coli"
/mol_type="genomic DNA"
/strain="O157:H7"
/sub_strain="RIMD 0509952"
/db_xref="taxon:562"
/names="pO157"
/note="RIMD 0509952 is a strain of enterohemorrhagic E.
coli, EHEC O157:H7"
join(92527..92721,1..2502)
/genes="tagA"
join(92527..92721,1..2502)
/genes="tagA"
/codon_start=1
/transl_table=11
/product="ToxR-regulated lipoprotein"
/protein_id="BAA31757.3"
/db_xref="GI:4666293"
/translation="MNTKMERWRTPMKLYLSCTILAPLAIGVFSATAADNNSAIYF
NTSOPINDLQGSAAVKAQSOILPAHPKEGDSQPHLTSLRKSLLVLRPVKADDTKTP
VOVEARDNNKILGTLTLPSPSLPDTIVHLQGVPEGGIDFTPHNCTKKIINTVAEVN
KLSDAGSSNKLHSLTNALVEIHTANGRVRIYLPQGPDLGKMWRFVSSAGYSSTV
FYGDRKVTLSVGNLTLLFKYVNGQWFRSGELENRRIITQAHIWSAELVIFGLNLV

IKQGNLSGLNDIKIGAPGELLHTIDICMLTTPRDRFPAKDKEAHREYFQTIPIVSR
MTVNNYAPLHLKEVMLPTGELLTDMDPGNGWHSGTMRQRIIGKELYSHGIDNANYGLN
STAGLCENSHYVYVVAQAAHNSRNGYANGIQVHGSGGGIVITLDSTLGNFSGHEVGH
NYGLGHYVDGFSVHRSANNSTWGDGKKRFIPNFVPSQTNKSCCLNNOQRPFP
DGHKPGDAMAGSGSPSAANRTMTYTPNSAIIQREFNKAVPDSKSTGFSKWNADT
QEMEPYEHTIDRAEQITASVNBLSKMAELMAEYAVVVKHMNGNWRNIYIPTASA
DNRGSLLTINHBAGNSYLFINGDERKVSQGYKFSVSDGQWKEKDDVVDREARKEQ
QFGVPTTIVGYPDEGLTSSYIYPAMYGAYGFTYSDSDSNLSDNDQQLQVDTKEQOL
PRFLANHRANTMNMKFINVPTESQPTQATLVCNKILDTKSLTPAPEGLTYTVNGQ
ALPAXENACITSVNSGKRYCLPVGORSYSLPDMITVGQEVYVDSGAKAKVLLSDWDN
LSYNRIGEFVGNVNPADMKKVKAWNGQIYLDFSKPRSMRVYVK"
/gene=".3464"
/gene="etpC"
2589..3464
/gene="etpC"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31758.1"
/db_xref="GI:3336999"
/translation="MLFPLSFRGDRGLFIKDIVLKMLTPNRLLCVILLIAGYOLVSVI
HFHWTQAAVSFGLSRVSAPEATVGDQTEBRFVFTLFGRASPLSSEGRAOETMPSLI
DOLLSGEDLDVRGILLYSSVAHSAI PAHNRQFSLSGEKVPSYDATISAI PDSHTIV
INYQKTVSLPLRYDNTERKADYNNLTVDVITQDNFRVSVFDIMSPSAVYVNT
LSGYRLIPGKHSSLFYNAGLHNDLAVSVNGSELRTDRAQIMKQLPELKEIKITVE
RDGQLYDAFIAVGEN"
3675..5432
/gene="etpD"
3675..5432
/gene="etpD"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31759.1"
/db_xref="GI:3337000"
/translation="MLNEQYTFQFPLSVLDVYGFVAVDMHNGILKVVRSKDAKTSVAP
VASDVSPTGDEVTVTRVPVSNVARDLAPLRQLNDNAGASVHYEESNVLLMTGR
AAVMKRLMEIVERDKVGNRSVATVPLTYASATDVARLVTELTKETDKTAIPAWMTAK
LVADERTNSLVSGEPIISOQRIISIIKQLDROEDVQGNKTVIYLYTAKARDLVVLTG
ISSIENDSKSPSTEALRKGVTIKSHQTNALILTGA PDVIRDLNENVSOLDIRRPQ
VIVEAIIARIQDADGLNIGIOMVNHKGAQVFTSLGPIITVMQTFQNEILSDQSNAL
LSMFNGIAGFYQGNWMLLTALSTSSKNDILATPSIVLTDNKEATFNVQGVPLVSG
SQTSGDNIFNTVERKTVGILKVRKPQINEGSVLLEIEQESVGVDATVATTTDIGA
TNTTRTVNMLVNGETVYVGGLLDKSRSESKVPLGIDIPVLGHLFRAKSEQTAK
RNLMFIRPTIIRERDGRFHAESAQYQSFNQVQSGKETTALTILNERQLRLSPQD
DTAFKVKAAIAFYAQEN"
5432..6937
/gene="etpE"
5432..6937
/gene="etpE"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31760.1"
/db_xref="GI:3337001"
/translation="MSRVQNVSESRPLPFSRQTIRILLREOQGNRVFCMEDTPPA
SALLERVRAEGLNVTTVTSAAFAEKQLVSSYQRDSDEARQWMAEIGNEMDFYVAGE
LPDRDLDLDDAPILRLINALETEAKESADITHIETVHRHLOVRFRIIDVGLVRIL
RLHRLNLSLISRIKVMARLDAEKRPQDGRMVLRIIGRAVDVRVSTLPSNHGERIV
LRLLDNVSLDLAALGNSQQNRHIDALIRRHPIGLIIVTGPFGSKSTLYLAALSIL
NPRDNIMTVEPVEYELDGISQTVNPKVDMTFARSLRAILRQDPDVLVGEIRGCE
TQAIVQAQSLTGLHVLSTLHTNSAAGLSRLMGIPPLFLLSTSLLAQLRVLRLC
PRCQPCQVSTELAMDMDIPPETTIWQAGCQMSPTGYHGRTGIHLELLIDDRIRTA
IYQGEGLGITRLAGSRYLTLRGDGRQKVLAGETSWEVVRVVTESRLQEEB"
6939..8162
/gene="etpF"
6939..8162
/gene="etpF"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31761.1"
/db_xref="GI:3337002"

gene

CDS

gene

CDS

gene

CDS

gene

CDS

/translation="MALPHYQASDIHCRKESGILEADSBARHARQLLEAOALJVRUDE
KQVHKHSLSILRFRPRGSSAEALALLTQRLATLVAAUSLPLEALDALLRQSEKPEQ
RNLIAAVTKVLEHSLAAAMFPGFPERLYCAMVAAGSETSLRDLVTLRSLADYTEQ
RQMRNRLQALLYPCVLTIVAVGVIAILLTAVPVKVEQFIHMKQTLPLSTRVLMGA
AEVQSTWGPVLLAAALGGIAGRMILHQPORLAFHLLRLPVGSRISRLNTARYA
RTLSILNASVPLLOAMHISGDVLSNDWARHQLATAELVREGVSLHQALEQTSLEPP
MWRMTASGENSGELDSMLERAADNDREFSTQMLALGLFEPFLVVGMAGVLFIVL
AIIQLPLLNNMMN"
gene 8193. .8627
/gene="etpG"
CDS 8193. .8627
/gene="etpG"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31762.1"
/db_xref="GI:3337003"
/translation="MRKHQHGFTLEIMVIVILGVLSLVVPLNMGKDKADROKV
MSDLVALESLDMVRLDNNRYPTTEQGLRALVSKPTVQPEPRNYRQDGYIRRLPQDPW
GGDYQLNPGQYSDIDIFSPGQGVNTEDDIGNWTLGNAQP"
gene 8624. .9178
/gene="etpH"
CDS 8624. .9178
/gene="etpH"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31763.1"
/db_xref="GI:3337004"
/translations="MSQRGFTILEMMLILLMGITAGLVLSPPDSQNHLOQORRL
QAOLDVALDRSQDGLLMGIQVRPDGKFKVLORGTAESPPTLAEGDIIWQVVMQTV
QPRAMGKLPDVRLEQLYRLGLQWMSHDDGAEPDILLPLPGGEVTFPRLFRQVG
EEAVGLQVDENGMLTFEGEVS"
gene 9193. .9540
/gene="etpI"
CDS 9193. .9540
/gene="etpI"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31764.1"
/db_xref="GI:3337005"
/translations="MTLLEIVIVLVFALGALMALQASTQORAGIGRMEKVLGWL
DNQVQLQLEKTPENGWGEKTSIFAGTEWYLRWGGPDSVDPQPSLEVEVRKKEET
SALVLSRSSVRE"
gene 9537. .10136
/gene="etpJ"
CDS 9537. .10136
/gene="etpJ"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31765.1"
/db_xref="GI:3337006"
/translations="MSQORVKFTLLEMLLAVFAALSISAFQVLQGGIRAHLSRD
KVORLRLQRCISQMERDLTQMLPRHSRGNELLLAAPHLLKSDDDGISFTNSWNP
AGMLPELQVGRYLRQKRLRSFYHVDHPGSPDVRVMDLVGHAFRLRFVFGND
WQARWDSGLPQAEVTEFLVMDDFAPLRLFLVSKETAE"
gene 10133. .11110
/gene="etpK"
CDS 10133. .11110
/gene="etpK"
/codon_start=1
/transl_table=11
/product="Type II secretion pathway related protein"
/protein_id="BAA31766.1"
/db_xref="GI:3337007"
/translations="MKUREQGVALLVLLLSLUMVTIAVIAERNORTFLRTVAQLDQ
LQKRWQGYTAETIAKQILQRSRSPKTHLAQNAQSERQFETRGDVRGQIVDAQ
CFNLNINAYGVDTITPYAARI FQQLLINLQVELLQARQVTAALRDWDIDKDPVRG
GADVEVMGMEPYLAANQPMQDSBLRIRIGDARLYRKLPLVPCVPTSDLSVNVN
TLSDAPLLAALFKPQDLSLPTVELLQRPRTGWSVAFLDAPLKDITSAAMPV
LAVSSNYFLVLRHVRGEHLFSQOTLMQWREPRRIIQRYQGLTMRVP"

gene 11473. .12321
Query Match 98.7%; Score 1166.2; DB 1; Length 92721;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 11 TGGAAAAAGCCAAATAAATAATTTGCCCATCCAGCGCGCTCCAGCTGAAAGTAGCGC 70
Db TGGAAAAAGCCAAATAAATAATTTGCCCATCCAGCGCGCTCCAGCTGAAAGTAGCGC 88880
Qy 71 TGTTCCTGTCGGGTATTTAAATGCAATTGACCGTCCCGGTATTTAAACAATGTGATAAATTA 130
Db TGTTCCTGTCGGGTATTTAAATGCAATTGACCGTCCCGGTATTTAAACAATGTGATAAATTA 88820
Qy 131 CTCGGTTACCGGAAAAACCGCTGAACAAATTCGGGCTGAAAAAGAGGATCCGCCGTATCT 190
Db CTCGGTTACCGGAAAAACCGCTGAACAAATTCGGGCTGAAAAAGAGGATCCGCCGTATCT 88760
Qy 191 GTTGCAATTTCCCTTAGCCCTGACTAGCCAGAGACAAATGATCTGTGCCGTCTCTGTTAAT 250
Db GTTGCAATTTCCCTTAGCCCTGACTAGCCAGAGACAAATGATCTGTGCCGTCTCTGTTAAT 88700
Qy 251 ATCAAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 310
Db ATCAAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 88640
Qy 311 TCGGGATAAAAAATTCGGCGAGTGGCGGCTCCATGAGACACATCCCCCAACGGGTAAACAG 370
Db TCGGGATAAAAAATTCGGCGAGTGGCGGCTCCATGAGACACATCCCCCAACGGGTAAACAG 88580
Qy 371 CGTCCCTGTACATTTCTTCTGAATGATCATCAGGAGTCCCGCCGCTCTCAGTCGGCGATAAC 430
Db CGTCCCTGTACATTTCTTCTGAATGATCATCAGGAGTCCCGCCGCTCTCAGTCGGCGATAAC 88520
Qy 431 GGGCAGCCGGAGACTGACGCTTTCAGCCAGTACATACCAACGCTTCATTTTCCGAAGG 490
Db GGGCAGCCGGAGACTGACGCTTTCAGCCAGTACATACCAACGCTTCATTTTCCGAAGG 88460
Qy 491 CATGACACACACTCGGCAATTCGGTAGACCGGTAAACGCTGGGAAAAAGGCACTGCCAT 550
Db CATGACACACACTCGGCAATTCGGTAGACCGGTAAACGCTGGGAAAAAGGCACTGCCAT 88400
Qy 551 TAACACATCTCCGCTCATTTCCAGAGTGTCTGCTGACGAGCGAGTCTTCGTTATTC 610
Db TAACACATCTCCGCTCATTTCCAGAGTGTCTGCTGACGAGCGAGTCTTCGTTATTC 88340
Qy 611 TTCACGCGCGGCCCCACCAACGAGCCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db TTCACGCGCGGCCCCACCAACGAGCCAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 88280
Qy 671 CAATACACGAGCATAAATTCATGTCTTTTTCGGGACGTAGCATCCCAACCTGAACGAT 730
Db CAATACACGAGCATAAATTCATGTCTTTTTCGGGACGTAGCATCCCAACCTGAACGAT 88220
Qy 731 AAGCGAAATTTGTCTGTGATGACGCCAGCGGTGATATGACAGGGGTAAACGGTCCGAT 790
Db AAGCGAAATTTGTCTGTGATGACGCCAGCGGTGATATGACAGGGGTAAACGGTCCGAT 88160
Qy 791 GGCTTCATTTATGCAATTCGGGGCAGTCCGAAACCCGGTGAATAACCGTTACCGGTGTCT 850
Db GGCTTCATTTATGCAATTCGGGGCAGTCCGAAACCCGGTGAATAACCGTTACCGGTGTCT 88100
Qy 851 GACACCTTCGGCATCAGATGCGCCATCATGCGGTGAGATAGGACAAACAATGAATCACA 910
Db GACACCTTCGGCATCAGATGCGCCATCATGCGGTGAGATAGGACAAACAATGAATCACA 88040
Qy 911 CAGATAATTCAGGAAAAAGTTCTGCTTACCGGTGATGTAGGTTTTTGTCTGACAAT 970
Db CAGATAATTCAGGAAAAAGTTCTGCTTACCGGTGATGTAGGTTTTTGTCTGACAAT 87980
Qy 971 AGTGAAGCGGTGACAGCATATCAGAGCGGTCTCAGTCTCTGCTATTTACTGTCTGCGCCACT 1030
Db AGTGAAGCGGTGACAGCATATCAGAGCGGTCTCAGTCTCTGCTATTTACTGTCTGCGCCACT 87920

QY 1031 ATGGCAGATCACAGATCAGGTTTAAATTTCCCGATAATCCGTCGAAGTCTGAGGATGGA 1090
Db 87919 ATGGCAGATCACAGATCAGGTTTAAATTTCCCGATAATCCGTCGAAGTCTGAGGATGGA 87860
QY 1091 AGGAAGTGAAGCTGTCTCTGAAAGAAATAAAGTGACATCATGCGCTTTTCTGGC 1150
Db 87859 AGGAAGTGAAGCTGTCTCTGAAAGAAATAAAGTGACATCATGCGCTTTTCTGGC 87800
QY 1151 TTCGGAGCAATTTTACTTTTCTCTGCAG 1181
Db 87799 TTCGGAGCAATTTTACTTTTCTCTGCAG 87769
RESULT 8
AX191725/c
LOCUS AX191725 92721 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 7 from Patent WO0149775.
ACCESSION AX191725
VERSION AX191725.1 GI:15209894
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1
AUTHORS Iversen, P.L.
TITLE Antisense antibacterial cell division composition and method
JOURNAL Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
source Location/Qualifiers
1. .92721
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"
ORIGIN
Query Match 98.7%; Score 1166.2; DB 6; Length 92721;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 11 TGAAGAAAAGCCAAATAAAAAATGCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 70
Db 88939 TGAAGAAAAGCCAAATAAAAAATGCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 88880
QY 71 TGTCTGTCCGTAATTAATGCAATGACCGTCCCGTATTAAACAAATGTGATAATTA 130
Db 88879 TGTCTGTCCGTAATTAATGCAATGACCGTCCCGTATTAAACAAATGTGATAATTA 88820
QY 131 CTCGGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCGTTATCT 190
Db 88819 CTCGGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCGTTATCT 88760
QY 191 GTTGCAATTTCCCTTAGCCTGACTAGCCAGAGACAAATGATCTGTCCGCTTCTGTTAAT 250
Db 88759 GTTGCAATTTCCCTTAGCCTGACTAGCCAGAGACAAATGATCTGTCCGCTTCTGTTAAT 88700
QY 251 ATCAACCCGCTACTCAATATCTTCTGTGGCTGGCTGCCATCATCGGAGGTTCCGG 310
Db 88699 ATCAACCCGCTACTCAATATCTTCTGTGGCTGGCTGCCATCATCGGAGCGTTCCGG 88640
QY 311 TCCGGATAAAAAATTCGGCAGTCCGCGGTCCATGAGACACATCCCCACCGGTTAACAG 370
Db 88639 TCCGGATAAAAAATTCGGCAGTCCGCGGTCCATGAGACACATCCCCACCGGTTAACAG 88580
QY 371 CGTCCCTGTCAATTTCTTGAATGACATCAGGGATCCCGCGCTCTCACTGGCGGATAAC 430
Db 88579 CGTCCCTGTCAATTTCTTGAATGACATCAGGGATCCCGCGCTCTCACTGGCGGATAAC 88520
QY 431 GGCGACCGCGGACACTGACGCTTCAGCCAGTACCATACAAACGCTTCAATTTCCGAAG 490
Db 88519 GGCGACCGCGGACACTGACGCTTCAGCCAGTACCATACAAACGCTTCAATTTCCGAAG 88460

QY 491 CATGACACACACACTGGCAATCCGGTAGACCGGTAAACGCTGGGAAAGGACACTGCCAT 550
Db 88459 CATGACACACACACTGGCAATCCGGTAGACCGGTAAACGCTGGGAAAGGACACTGCCAT 88400
QY 551 TAAACATCTCGCTCATTTCCAGGTTCTGTCTGCTGACGAGAGGTTGCTTGTGTTTC 610
Db 88399 TAAACATCTCGCTCATTTCCAGGTTCTGTCTGCTGACGAGAGGTTGCTTGTGTTTC 88340
QY 611 TTCAGCCCGCGGCCACACAGAGCCAGCAATGATTTCCCTTCCATCTTTCAGCTGATA 670
Db 88339 TTCAGCCCGCGGCCACACAGAGCCAGCAATGATTTCCCTTCCATCTTTCAGCTGATA 88280
QY 671 CAATACACGAGCATAAATTTCTGCTTTTCGGGACGTAGCATCCCACTGAAACGAT 730
Db 88279 CAATACACGAGCATAAATTTCTGCTTTTCGGGACGTAGCATCCCACTGAAACGAT 88220
QY 731 AAGCGAACAATTTCTGCTGATGACGCCAGCGGTGGATATGACGGGGTAAACGCTGCGAT 790
Db 88219 AAGCGAACAATTTCTGCTGATGACGCCAGCGGTGGATATGACGGGGTAAACGCTGCGAT 88160
QY 791 GGCCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGTCCT 850
Db 88159 GGCCTTCATTATGCAATGCGGGCCAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGTCCT 88100
QY 851 GACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGGACAAACAAATGAAATCACA 910
Db 88099 GACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGGACAAACAAATGAAATCACA 88040
QY 911 CAGATAATTCAGGAAAACGTTCTGTGCTTACGGGTGATGTAGGTTTGTCTGACAAT 970
Db 88039 CAGATAATTCAGGAAAACGTTCTGTGCTTACGGGTGATGTAGGTTTGTCTGACAAT 87980
QY 971 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCGCTATATATTAATCTGTCAGGCCACT 1030
Db 87979 AGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCGCTATATATTAATCTGTCAGGCCACT 87920
QY 1031 ATGGCAGATGACAGATCAGGTTTAAATTTCCCGATAATCCGTCGAAGTCTGAGGATGGA 1090
Db 87919 ATGGCAGATGACAGATCAGGTTTAAATTTCCCGATAATCCGTCGAAGTCTGAGGATGGA 87860
QY 1091 AGGAAGTGAAGCTGTCTCTGAAAGAAATAAAGTGACATCATGCGCTTCTTCTGGC 1150
Db 87859 AGGAAGTGAAGCTGTCTCTGAAAGAAATAAAGTGACATCATGCGCTTCTTCTGGC 87800
QY 1151 TTCGGAGCAATTTTACTTTTCTCTGCAG 1181
Db 87799 TTCGGAGCAATTTTACTTTTCTCTGCAG 87769
RESULT 9
AF043470
LOCUS AF043470 5612 bp DNA linear BCT 25-JUN-1998
DEFINITION Escherichia coli plasmid p0157 ecf4 gene, partial cds; and ecf3, ecf2, and ecf1 genes, complete cds.
ACCESSION AF043470
VERSION AF043470.1 GI:3253288
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE 1 (bases 1 to 5612)
AUTHORS Boerlin, P., Chen, S., Colbourne, J.K., Johnson, R., De Grandis, S. and Gyles, C.
TITLE Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids and the locus for enterocyte effacement in shiga toxin-producing E. coli
JOURNAL Infect. Immun. 66 (6), 2553-2561 (1998)
MEDLINE 98261495
PUBMED 9596716
REFERENCE 2 (bases 1 to 5612)
AUTHORS Boerlin, P. and Gyles, C.
TITLE Direct Submission

JOURNAL	Submitted (17-JAN-1998) Pathobiology, ON Veterinary College, University of Guelph, Guelph, ON N1G 2W1, Canada
FEATURES	Location/Qualifiers
source	1. .5612 /organism="Escherichia coli" /mol_type="genomic DNA" /strain="4304-FT14" /serotype="O157:H7" /db_xref="taxon:562" /plasmid="pO157" /note="produces Shiga toxin" complement(<1. .705) /gene="ecf4" complement(<1. .705) /gene="ecf4" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24348.1" /db_xref="GI:3253289" /translation="MVPPSAVLCHNEISHQIPVNMKNIRTEIPFNLTLCPFRYYM TWGTGIIICVFMVPPALRDLGLKGLMGLRGLKSRQALINLSLCPPEYDKEKE NIVDAMFATASMAVLAELALSGDPDKISHRIWNGLIEI VERMAQNEKVI FLVPHAW QYDI PAMLMASAGRKMAAFHNPQRPVVDVYVNSVRRRFGKGLHARDGIIASFVRSVR QYGYIYLPDQDHGPEF" complement(751. .2499) /gene="ecf3" complement(751. .2499) /gene="ecf3" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24349.1" /db_xref="GI:3253290" /translation="MHLNTGONRPTFSNALGWALFYFGFSTLLQVLIFFSSGSGTN GIRDSLLFSCWLIPVFLYIPDKIILAAVVGFIIMGTSILAALCTYLIYGHFPSQVIF VNFETNAEAGEYFSQPSLKLILISLYTVASVFLWLRUPVYIPLWRRIVSFLIL YALLHPVLKSLRQEPFLNDLTKLASRMPEAPAPQWOFVSSYTYHQQLMALTFPLNE NSALPGLNLDSEGERPTLVLVIGESTORERMSLYGLRETTPELDALRKTDPLGT VFNVVASRPVYIHALQALTFANEKNPDLYLTQPSLMMNMKQAGYKTFWITNQQTIT ARNTMLTVFSRQTDROYQNMQRQTSAREYDVTNLKPPREVLDNAPAKKLIIVHLGRT HIKYKIRPEQGRFPGDGLTGHIPGLNAKELEVINDYDANLFDNDHVAASLIDKPRAT APDGLLYFSDHGEVYDTPPYKQGRNEDNPRPMYTVPFLLWSEKWHAAHPRDFS QYDVKYLSABLHITMSDLAGLTGYDPTPSLVNPNQPRETTRWIGNPYKNGLTDFD LFRMVSHIFLSSVPF" complement(2589. .3695) /gene="ecf2" complement(2589. .3695) /gene="ecf2" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24350.1" /db_xref="GI:3253291" /translation="MMKILTFESSSDIGGOELQALQMTALQKHGSVILLACREKSKI APEARKGHDVTFFIPNLSHLPLFRRIIEFKPDPLVICHSGHDSNIAGLSRLIC CHRFGIYKQKVIYKRTKTFSLNLYCDFIVVPSAMMAHMAEGVTPVTVIIPGDFD PALHNEARPLPLHIAWAASADNVPLIVQWMLRPEKGFHMLVLYOLKMEGSKFR WLVGAGREEEASLRQOTHELGMSDVL MAGALFPALPVYRIASVVVPSSENAEAFGM VLAESVSGVPVIASETGIPDVIQKNVTGLPLPVGDVSAWTAUGALRDFLSRPERFRM AASAREDEYRFDINRTAQIIVSLASQAKGKCNR" complement(3695. .4516) /gene="ecf1" complement(3695. .4516) /gene="ecf1" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AAC24351.1" /db_xref="GI:3253292" /translation="MLNARHLPLVMYHHVSRCPGLVTLSPETCEQMKWLADNNWRTV TSALEYFYQGGTFLPRKSVMLTDFDGYLDNWLRAWPVLYKELNHLHAFILITGLIGKN VRSRQEQYSHRDCERLIADNRWSVEVRMDSGLVFFHVHTHSHKRWDRLS
gene	
CDS	
gene	
CDS	
gene	
CDS	

ALONGHADSIRMFSLINKMLSSQIKELLKVKHGLFMAJONGHTKAIMAYGDILKI
LPQHEVIDELWLNKNGNSGLFMAYFNGHTETIRAFPCNLIKKNYSFTIRLVEMLSA
TKNGDIGFVSVNVRDKETILEYRIIKENNELPDITAEQFSKMKKTIEIINRFN
HFL"
complement (11642..12361)
/gene="ospP"
complement (11642..12361)
/gene="ospP"
/note="ORF12, length = 239 aa, 63.3% id in 245 aa overlap
with STMKFA_1 (Salmonella typhimurium virulence plasmid
Mkfa = 241 aa)"
/codon_start=1
/transl_table=11
/product="ospP, secreted by the Mxi-Spa secretion
machinery, function unknown"
/protein_id="CAC05773.1"
/db_xref="GI:12329042"
/db_xref="GOA:Q99Q87"
/db_xref="UniProt/TREMBL:Q99Q87"
/translation="MPKIKCKLKLNSLVVVKSEIQMLSANERLKNFNILYNQIR
QYPAYIPKVASNPVTSYDIOFFSVMTYQGFQVNHSGDVFIHACRENPQSGDFVGDK
FHSIAREQVPLAFQIISGLFLFSDSPIDKWKITDMNRVQSQSRVGIQAQFTLYVKS
QESQYQALLIKIRQITMCLNLSRLSKTAPGEYPASVDPEDWKVSVYRNEILSDR
DGSROSLREBPYRLMIE"
complement (13393..13971)
/note="unnamed protein product; ORF13, length= 192 aa,
unknown"
/codon_start=1
/transl_table=11
/protein_id="CAC05774.1"
/db_xref="GI:12329043"
/db_xref="GOA:Q99Q33"
/db_xref="UniProt/TREMBL:Q99Q33"
/translation="MKVFSKSLGYIPHDYINKKHTIDFNVDVVRKAVLSKINELNAC
HKVAIPFAEKONEITKKDKAKIIDLTFENTSIIEFQKLNISERTLNSLYITGESGF
VSFVNRGKICHTAYVKSQSDNSMTYHANGSSIDKYITDMCLICMRHIESGTGIIFW
LDEKLSAIAEFNWEKWRAPCSAKNLYKCV"
14014..14241
/note="IS2.02, 90% id over 228 nt with IS2, from 4 to 230"
complement (14367..14629)
/note="IS21.01, 97% id over 263 nt with IS21, from 1825 to
2098"
complement (14630..15558)
/note="IS91.01, 97% id over 929 nt with IS91, from 191 to
0"
15559..15924
/note="ISShf5.02a, id to ISShf5 from 1 to 366"
complement (15925..17350)
/note="IS4.01, 99% id over 1426 nt with IS4, from 1 to 0"
17351..19039
/note="ISShf5.02b, id over 1689 nt with ISShf5, from 335
to 0"
19326..19523
/note="IS3.01, 88% id over 198 nt with IS3, from 1062 to
0"
20459..20601
/note="IS630.02, 84% id over 143 nt with IS630, from 999
to 1141"
20964..21641
/gene="ospD1"
20964..21641
/gene="ospD1"
/note="ORF21, length = 225 aa, 34.3% id in 140 aa overlap
with ECHSET2B.1 (S.flexneri SenA = 565 aa)"
/codon_start=1
/transl_table=11
/product="OspD1, secreted by the Mxi-Spa secretion
machinery, function unknown"
/protein_id="CAC05775.1"
/db_xref="GI:12329044"
/db_xref="GOA:Q99Q96"
/db_xref="UniProt/TREMBL:Q99Q96"

Query Match 27.8%; Score 328.6; DB 1; Length 213494;
Best Local Similarity 59.3%; Pred. No. 7.1e-95;
Matches 578; Conservative 0; Mismatches 394; Indels 3; Gaps 1;
Qy 210 TGACTAGCAGAGACACAATGATCTGTGCGGTTCTGTTAATATCAAAACCGGTACTCAATA 269
Db 187295 TGCTTCGCTAATGTGAGTATTTTAAACGAGTTTTTATTAATATCAACCGCTCTTCTATA 187236
Qy 270 TCTTCTCTGGCGTGGCTGCCATCATCCGGAAGCGTTCCGGTCCGGGATAAAAAATTCGGC 329
Db 187235 TCCTGTTTTGCTGGCGAGCCATCTGATAAAAAACGCCAGGGTCATTAAAAAAATCATTC 187176
Qy 330 AGTGGCGCGGTCCATGACAGACATCCCCACCGGTAACAGCGTCCCTGTGCACATTTCTC 389
Db 187175 AGGGGCGACATCAATGCGTCTTATTAACCTGCTGTGTAACATGTGCCGGTCTGGTTATTC 187116
Qy 390 TGAATGACATCAGGGATCCCGCGCTCTCACTGGCGATTAACGGGACACGCCGAGACTGAC 449
Db 187115 TGAATAACATCAGGATTTCCACCAATCTGACTGGCCAGTACAGGCACAGAAAATGCCGAT 187056
Qy 450 GTTTCAGCCAGTACCATACCAAAACGGTTCTATTTCCGAAGGCATGACACACACATGGCA 509
Db 187055 GCTTCTGCCAGTACCATACCAACGATTCGTTTTCTGAAGGCAGAACCCAGACTGGCA 186996
Qy 510 ATCGGTAGACCGGTAAACGCTGGGAAAGGGACCTGCCATTAACACATCTCCGCTCAAT 569
Db 186995 ACCCGATATACGGGGCGCGCAGGAAAAACAATTGTACGAAATAAAAAATCATATGCGATG 186936
Qy 570 CCCAGGTGTTCTGTCTGCTGACGACGAGTCTTCGTTATTTCTTCAACCCCGCGGCCACCC 629
Db 186935 CCCATCTGTAATCTGATACTGTAAATGCTCCGGCAGTTTCAGGCGAACAGACCCACACA 186876
Qy 630 ACAGGCGAGCGAAATGATTTCCCTTCATCTTCAGCTGTGATCAATAACACGACGATAAAT 689
Db 186875 ATCAACCAACAGAACTGTCTGTCATTTTAAATGGAACAGTAAATTCACATAAAT 186816
Qy 690 TCATGTCTTTTTCGGGACGTAGCATCCCACTGAACGATTAACGGGAAACATGTCCTGCT 749
Db 186815 TCGTGCCCTTTTTCGGGGCGCAACATTCCTACCTGAGCAATTAACGGGGGCAATCCCTTCGG 186756
Qy 750 GATGACGCCAGCGGTGGATGATGACGGGGTAAACGGTCCGATGCTTCATTTATGCAATGCG 809
Db 186755 GAGCCAGCCCAAGAAAGAACATTTGGAGGCAACTGTTTCGGGNATCAACGTATATTTTC 186696
Qy 810 GGCAGTTCGAAACCCGGTGGAAATAACGGTTTACCGGTGCTCTGACACCTTCGCGCATCAGA 869
Db 186695 TCGAAATCAAAAGCTTCGAGGCACACAGTAACCGGGTTCGACATCCTTCTGCTCCAGA 186636
Qy 870 TCGCCCATCATGGGTGAGATAGGCACACAAATGAAATCAACAGATTAATTCAGGGAAAC 929
Db 186635 TGTGTCTTCATCTTGTTCGGGAAACAATCACCTCATCGAAAAATGATTTATTGAAAAA 186576
Qy 930 GTTCTGTCTTACGGGTGATGTAGGTTTTTGTCTGACAAATAGTGAAGCGGTG--ACAG 986
Db 186575 ACTTTTGTCTGCTCAATATGCTTTTGGCTGATATTTCTGATGATGATGTTTCTTA 186516
Qy 987 CATATCAGACGGCTCAGTCTCTGTATATTAATCTGTCAATGGCCACTATGCGCAGATGAC 1046
Db 186515 GTGAATAACCGTACTAAACCAACAATATTCATATGCTATGCTGACAAACGATTGCA 186456
Qy 1047 TCAGGTTTAAATTCCTCGGATAATCCGTCGAAGTCTGAGGATGGAAGAAAGGTGAGCGTG 1106
Db 186455 TTGGGCTGAAAACCATGAACATATTCCGAGTAATCTCCATCAGTAGGATGTTGAGAGACTG 186396
Qy 1107 TTCCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTCTGCTCTCCGAGCAATTTTA 1166
Db 186395 TTTCGAAATAACCGGATGTGATATCAATCCCAATTTACTGCTTCAAAAGCAATTTTG 186336
Qy 1167 CTTTTTCTCTGCAG 1181
Db 186335 CTGTTTTCCCTGCAG 186321

/translation="MWVSIAIASTHLVYVIRTEYTYGTRRLQTELDANGNIIVGRDLRALQ
LRKELRLKCKQKRRFRATNSDHLNLPVNLNQNFPTTAPNQVWADSVVQAFRQNP
TEGAGRETAAYVR"
gene
complement (6107. .6925)
/genes="CP0008"
CDS
complement (6107. .6925)
/genes="CP0008"
/notes="similar to a 272 aa protein from Shigella flexneri
pir: T44484"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL72317.1"
/db_xref="GI:18462625"
/translation="MCQVFGSRSGSYNNVQHEPSDEKQSDERLKLKLEIKVAHIRTRET
YGTRRLQTELAENGIIVGRDLRALRKLRLCKQKRRATNPNHNLVAPNLNLQ
TFAPTAPNQWADLVYATQGEKLYLAGIKDVYTCIYVAGMERMTKELTGKALFM
ALRSQRPPLHGLSDRSGSCAYDYRVIQESGLGKTSMSRKGNCYNDMELTSPFWGLT
KNESLHYFNNRDEAISVREIYEIFNRRHSRLGNISPAAFREKYHQMAA"
gene
complement (7435. .9153)
/genes="ospD2"
/notes="synonym: CP0009"
CDS
complement (7435. .9153)
/genes="ospD2"
/notes="similar to a 569 aa protein from Shigella flexneri
emb: CAC05772.1"
/codon_start=1
/transl_table=11
/product="OspD2"
/protein_id="AAL72317.1"
/db_xref="GI:18462545"
/translation="MSSMPLNKLTFSSIFSTKNSLSDMSVNRDNRITITSSIMRVNS
SELIQKNKTAPYSEKRNNEVNINGAKDIYGRQIVCRHLASVWNNFMETNGKYNV
QLLSHPALAKNCLTEDEPSPKSPAVIYFVNKKGTVI TNFYNMKNGDPVRLS
ACTLNHOMALGKTKIQRVESEKWWQFDPNRTVTHKRTVPTCDSPHELSQLSADFF
DDFYMKIYGLPQGOVAFEDRHSNPLTNFKLLPDELINSEVYHAIKNTLSEVLIL
MEKYANEISQSLNLLATRSSDGTAPYALQNGSDIIQVYKILNMCNLSQSTI
LTLAAVANNVPGLCWSFMNGVHDTIKAYGEIVFKTSLDRLYLLAAKDSHLPG
LFPALQNGHADSIRFMGLSNLKMLESSEQIKLLKVKHGLFMALQNGHTKAIMAYGDI
LKLIIPHQEYIDELLKIKNPNGTSGLFMAFYNGHTETIRAFCNILKNYSPTTIRLIVEM
LSATNKGDIQGVFVSVNRDKETILEYCRIIKENNLPDPTIAEQFSKMKMKTPIEIN
RNFHFL"
gene
complement (9576. .10295)
/genes="ospf"
/notes="synonym: CP0010"
CDS
complement (9576. .10295)
/genes="ospf"
/notes="similar to a 239 aa protein from Shigella flexneri
emb: CAC05773.1"
/codon_start=1
/transl_table=11
/product="Ospf"
/protein_id="AAL72315.1"
/db_xref="GI:18462543"
/translation="MPDKKPCLLNLDLSNVVRSEIPQMLSANERLKNFNILYQIR
OYAPYIPKVASNPVTYDII COSFSVMYOGFOIVNHSGDVEIHACRKNPOSKGVFDGDI

	Query Match	27.8%	Score 328.6;	DB 1;	Length 221618;
	Best Local Similarity	59.3%	Pred. No. 7.2e-95;		
	Matches 578;	Conservative 0;	Mismatches 394;	Indels 3;	Gaps 1;
Qy	210	TGACTAGCCGAGACACAATCATCTGTGCCCTTCGTTAAATATCAAAACGGGTACTCAATA	269		
Db	193635	TGCTTCGCTAATGTGAGTATTTTAAACGACGTTTTATTAAATATCAAAACGGCTCTTCTATA	193576		
Qy	270	TCATTCTGTGGCGCTGGGTGCCCATCTCGGAAGCGTTCGGGTCCGGATAAAAAATCGGCG	329		
Db	193575	TCCTGTTTTCCTCGCGAGCCATCTGATAAAAACGCCAGGTCATTAAAAAAATCATTC	193516		
Qy	330	AGTCCGCGGTCATGACAGACATCCCCCAACGGGTAAACAGGTCCTGTGCATATCTTC	389		
Db	193515	AGGCGGCACATCCATGGGTGCTTAACTGCTGTGTAAACATGTCCGGCTGTGTTATTC	193456		

Qy	390	TGAATGACATCAGGATCCCGCCCGTCTCACTGGCGATAACGGGCACGCCGAGACTGAC	449
Db	193455	TGAATAAACATCAGGATTTCCACCAATCTGACTGGCCACTACAGGCACAGAAATATGCCGAT	193396
Qy	450	GCTTCAGCCAGTACCATACCAAAAGCTTCATTTTCCGAAGGATCACCACACACTGGCA	509
Db	193395	GCTTCTGCCAGTACCATACCAAAAGATTCGTTTCTGAAAGCGAAGAACCAACGACTGGCA	193336
Qy	510	ATCCGGTAGACCGGTAAACGCTGGGAAAAAGGCACTGGCAATPAAACATCTCCGTCATT	569
Db	193335	ACCCGATATACGGGGCGCAGGAAAAACATTTGTGAGCAATAAAAAACATCATCATGCAATG	193276
Qy	570	CCCAAGTGTCTGTCTGCTGACGAGCGTCTTCGTATTCTTTCACGCCGCGGCCACCC	629
Db	193275	CCCATACTGTCAATCTGATACGTGTAAATGTCTCCCGCAGTTTCAGGCGAAACGACGCCCA	193216
Qy	630	ACGAGCCAGCGAAATGATTTCCCTTCCATCTTTCAGCTGATACAAATACAGCGACGATAAAT	689
Db	193215	ATCAACCAACAGAACTGTGCTCCATTCATTTTAAATGGAACAGTAAATTTCAACATAAAT	193156
Qy	690	TCATGTCTTTTTCGGGACGTAGCATCCCCACCTGGAACGATAGCGGAACATTTGCTGCT	749
Db	193155	TCGTGCCCTTTTTCGGGCGCAACATCTCTACCTGAGCAATAAACGGGGCATCCCTCTCGG	193096
Qy	750	GATCAGGCCAGGCGTGGATATGCAAGGGGTAAACGGTTCGATGGCTTCATTATGCAATGCG	809
Db	193095	GACGCCAGCCAAGAAAGAACATTTGGAGGCAAACTGTTTCGGAATCAACGTATTAATTTTC	193036
Qy	810	GGCCAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGCTCTGACACCTTTCGCCCATCAGA	869
Db	193035	TGGAAATCAAAGCCTGGAGGCACAACAGTAAACCCGGTTCGACATCCTTCTGCTCCAGA	192976
Qy	870	TGCGGCATCATGGGTGAGTAGGCAACAATGAATCACACAGATAATTCAGGGAANAAC	929
Db	192975	TGTGCTTTCATACTGTTTCGGGAACCAATCACTCATCGCAAAAATGATTTATTGAAAAA	192916
Qy	930	GTTCTGTGCTTACGGGTGATCTAGGTTTTTTGTCTGACAAATAGTCAAGCGGTG---ACAG	986
Db	192915	ACTTTTGTTTTTTCGTGTCMAATATGTCTTTTGCCTGATTAATTCGAAATGGATGTTTCCTA	192856
Qy	987	CATATCAGACGGCTCAGTCTCGCTATAATTACTGTTCATGGCCCACTATGGCAGATGCCAGA	1046
Db	192855	GTGAATAACCGTACTAAACCAACAATAATTGCTATCATGCCACTGTGACAAACGATTGCA	192796
Qy	1047	TCAGGTTTAAATTTCCCGATTAATCCGTCGAAGTCTGAGGATGGGAAGGAAGGTGAAGGCTG	1106
Db	192795	TTGGGCTGAAACCAACCATGAACATATTCCCGAGTAATCTCCATGCAGTAGGATGTGAAGACTG	192736
Qy	1107	TTCTCTGAAAGGAAATAAAGTGACATCATGCCCTCTTTTCTTCTGGCTTTCGGAGCAATTTTA	1166
Db	192735	TTTCGAAATAACGGAATGTGATATCAATCCCCAATTTACTGGCTTCAAAGCAATTTTG	192676
Qy	1167	CTTTTTTCTCTGACAG	1181
Db	192675	CTGTTTTCCCTGCAG	192661

RESULT 12
AF348706/c
LOCUS
DEFINITION AF348706 221851 bp DNA circular BCT 24-APR-2001
Shigella flexneri large virulence plasmid pWR501, complete plasmid
sequence.
ACCESSION AF348706
VERSION AF348706
KEYWORDS .
SOURCE AF348706.1 GI:13310487
ORGANISM Shigella flexneri
Shigella flexneri
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Shigella.
REFERENCE 1 (bases 1 to 221851)
Venkatesan,M.M., Goldberg,M.B., Rose,D.J., Grotbeck,E.J.,
Burland,V. and Blattner,F.R.
AUTHORS Complete DNA sequence and analysis of the large virulence plasmid


```
CDS
/gene="ORF186"
763..1851
/gene="ORF186"
/codon_start=1
/transl_table=11
/product="ORF186"
/protein_id="AAP79030.1"
/db_xref="GI:32307067"
/translation="MNILFTESPNIQGQELQAVQMKALKMGHSLVLCRENSKIA
FEASKGIDITFALFNSLHPTAWRLGLVHGFQPNIAIVCHSGDSNIVGLRFLPTW
KHPFRIROKTYLTQTKVFSINHFCDEVLVPGTSMKTHLEQSGCRTRVTVVPGDF
OKLYVDSRLSPKTVLSWLASRGCPIVQGVKMLRPEKHEFMNLFLFLKMGROFC
WLVGSGSELRHLQYQIDSMGHDDVFIAIDNVFPAPVIRVASLVLPSENESFGM
VLAASAFVSVLASIGQIPDVQNNQGTLLPAGNKHAMWCALNDFNDFPGRFYQM
AQAKQIDIERFDINKILTLAKHK"
gene
1856..2806
/gene="virk"
1856..2806
/gene="virk"
/notes="required for proper localization of IcsA (VirG) at
the surface of bacteria"
/codon_start=1
/transl_table=11
/product="virk"
/protein_id="AAP79031.1"
/db_xref="GI:32307068"
/translation="MFSVSNLSFIFGLKRVFSSDSLPGKWEHRKFRPMYILRCAINP
VASIRVYELRSLPCEDILAIHPTLPARIHRYPLHKGRAWTRGOYILHVRVFNOL
PEKYSKFLPQKSVLSVQFVIGKNGENFDIQCSSGDFREGELMLSLFYNKTVITLTF
SVILITQNGHAFIGGLQAPKNTGPDVIRCATRACYGIFPKRLIIFAPCALMKACNIS
ECLAVSEHSHVFLQRLWYQKRTFVAVSDFWSEVAGTKCGDWYRLPTQVIRKPLSD
IASKRSEYRKRYALLDYIHTETIRSLDAYPVHSEHODLN"
2871..3815
/gene="mabB2"
2871..3815
/gene="mabB2"
/notes="probable lipid A transacetylase"
/codon_start=1
/transl_table=11
/product="MabB2"
/protein_id="AAP79032.1"
/db_xref="GI:32307069"
/translation="MKKIKSEFIPFKKNYLSPVYWSWTFLLGMIAGISNFPSPRDP
VLAKTGRWAGRLSKARRRATINLSLCPFKSDTEREIIIVDKHFATALQSIYMAELA
IRGPEKQFVRVFWKGLIEELIRYNNRNVIPLVPHGWSVDIIPAMLIQAQEKMAAMFH
QORNPVIDVWNSVRKFGRLHSRDKIPFIQSVROGVWGYLIPDQDHGPEYSEFA
DFEATYKATLPIGLRLNLSQAMILPLFVYDEKHFLEIYVRPPMDACIASADNKM
ARQMKTVIILVGSHPSEQIYIWLKLLKTRKSNEADPY"
ORIGIN
```

```
Query Match 27.7%; Score 327; DB 1; Length 3815;
Best Local Similarity 59.2%; Pred. No. 9.8e-95;
Matches 577; Conservative 0; Mismatches 395; Indels 3; Gaps 1;

QY 210 TGACTAGCAGAGACAATGATCTGTGGCGTCTCTTAAATATCAAAACGGTACTCAATA 269
DB 1844 TGCTTCGCTAAATGTGAGTATTTTAAACGCGAGTTTATTAATATCAAAACGGCTCTTCTATA 1785

QY 270 TCTTCTCTGGCGTGGCTGCCATCATCCGGAACGTTCCGGTCGGGATAAAAATCGCG 329
DB 1784 TCCTGTTTTGCTCGGCGAGCATCTGATAAAAACGCCCGAGGTCATATAAAAATCATTC 1725

QY 330 AGTGGCCCGGTCCATGCAGACACATCCCCACCGGTAACAGCGTCCCTGTCTCATATCTTTC 389
DB 1724 AGGGCGCACATCATCGGTGCTTATTACCTGTGTTAAACATGTCCCGTCTGGTTGTTTC 1665

QY 390 TGAATGACATCAGGATCCGCCCGTCTCATCTGGCGATTAACGGGACGCGCGGAGACTGAC 449
DB 1664 TGAATAACATCAGGATTCACCAATCTGACTGGCCAGTACAGGCACAGAAAATGCCGAT 1605

QY 450 GCTTCAGCCAGTACCATCAACACGCTTCATTTCCGAAGGATGACACCCACCTGGCA 509
DB 1604 GCTTCTGCCAGTACCATACCAACAGTATCGTTTTCTGAAGGCGAACCACGACTGGCA 1545
```

RESULT 14
AF134403/c
LOCUS
DEFINITION
Escherichia coli plasmid pAA2 Shf (shf), hexosyltransferase homolog (capu), and VirK (virK) genes, complete cds.
ACCESSION
AF134403
VERSION
AF134403.1 GI:4959583
KEYWORDS
Escherichia coli
Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.
REFERENCE
1 (bases 1 to 3500)
AUTHORS
Czezuln,J.R., Whittam,T.S., Henderson,I.R., Navarro-Garcia,F. and Nataro,J.P.
TITLE
Phylogenetic analysis of enteroaggregative and diffusely adherent Escherichia coli
JOURNAL
Infect. Immun. 67 (6), 2692-2699 (1999)
MEDLINE
99270924
PUBMED
10338471
REFERENCE
2 (bases 1 to 3500)
AUTHORS
Czezuln,J.R., Whittam,T.R., Henderson,I.R. and Nataro,J.P.
TITLE
Direct Submission

JOURNAL Submitted (11-MAR-1999) Vaccine Development, Univ. of Maryland, 685

W. Baltimore St., Baltimore, MD 21201, USA

FEATURES

source

1. 3500

/organism="Escherichia coli"

/mol_type="genomic DNA"

/strain="042"

/db_xref="taxon:562"

/plasmid="pAA2"

/note="enteroaggregative"

558..1400

/gene="shf"

558..1400

/gene="shf"

/note="hypothetical; similar to Shigella flexneri Shf"

/codon_start=1

/transl_table=11

/product="Shf"

/protein_id="AAD34404.1"

/db_xref="GI:4959584"

/translation="MLNGGILFKANHVPLVHVHSHCPGLVLTSPETFRKQMKWLA

ENMKTLSSDELFFYRGGKLPKRSVNLTFDDGYLDNWFQVYPLNFPNLKAHVFLIT

SFTGNGPVRHSPOKYSHRDCEHQIATGNADNVMRWSEVNEMLQSLGVLFVHTYTH

TRWKKTSREOCKLRDLGLSREYLVKWTCKSKHLCPREGYNKYDIQIABELG

FHYLYTTERMNPAKGAARIGRISTKERESCAWLKRLFYTYTTFPSSLLALYKQPR

LPDD"

1403..2224

/gene="capU"

1403..2224

/gene="capU"

/note="CapU; hypothetical"

/codon_start=1

/transl_table=11

/product="hexosyltransferase homolog"

/protein_id="AAD34405.1"

/db_xref="GI:4959585"

/translation="MNLPTSEPNIGQELQAVQMKALKMGHVSLLVCRENSKIA

FEASKGIDFTALFRNSLHPTVWRLLGLVHSFQDAIVCHSGHDSNIVGLRVFTW

KHPFRIQRTKLTFRKFSINHFDEIVPGTSMKTHLEQEGCRVTVVPPGDF

QELVDSRSLPSVLSWASRRGCPVIAQGMRLPEKGGDPLCMLIVGFLPELREH

LOQIDSMGMDHVDFTADNVLPVPPVIGLPMVWFCLQKTNLLVWCQKRRHFLCL"

2453..3403

/gene="virk"

2453..3403

/gene="virk"

/note="similar to Shigella flexneri VirK"

/codon_start=1

/transl_table=11

/product="VirK"

/protein_id="AAD34406.1"

/db_xref="GI:4959586"

/translation="MFSISNLSIFGLKRVFSDSLPGKWEHRKFRPMYILRCINP

VSVIRYVYELRSLPDIILAIHPTLPARIHPRYLKGGRAWTRGOVILHEHYRFVONL

PEKYSKFLPPOKSVSLVQFQKNGENFDIQSPSGFDREGELMLSLFPYNTVIARLT

SVLTQNGHAFITGGIGAPKNTGPDVIRCATPACYPGLPKRIIFPAFCALMKACNVS

ECLAVSEHSVFLQKTKTQKRTFVAVSDFDWDSVFGTKGDMYKLPQTVIRKPLTD

IPIKRSYRKQYALLDYIHETTRISLDAYPVHSHQNSN"

ORIGIN

Query Match 20.9%; Score 246.6; DB 1; Length 3500;

Best Local Similarity 56.4%; Pred. No. 2.1e-68;

Matches 550; Conservative 0; Mismatches 379; Indels 46; Gaps 3;

QY 210 TGACTAGCCAGACACAAATGATCTGCGCGTCTCTGTTAATATCAACCGGTACTCAATA 269

DB 2441 TGCTTCTGCTAATGTAGATATTTTAAACACAGTTTATTATATCAACCGCTCTCTATA 2382

QY 270 TCTTCTTGGCGTGGCTGCATCATCCGGAAGCGTTCCGGTCGGGATAAAAATCGCGC 329

DB 2381 TCCTGTTTTGCCAGGCGAGCATCTGATAAAAACGCCAGGGTCATTAAAAAATCATTC 2322

QY 330 AGTGCCCGGTCCTCAGACACATCCCCACGGTAACAGCGTCCCTGTACATCTTTC 389

SHFVIRK Shigella flexneri plasmid pMYSH6000 virulence protein (virk) gene, complete cds.

LOCUS D11025 1642 bp DNA linear BCT 03-FEB-1999

DEFINITION D11025 1 GI:216943

VERSION D11025 1 GI:216943

KEYWORDS virulence protein.

SOURCE Shigella flexneri

ORGANISM Shigella flexneri

REFERENCE 1 (bases 1 to 1642)

AUTHORS Nakata,N., Sasakawa,C., Okada,N., Tobe,T., Fukuda,I., Suzuki,T.,

TITLE Komatsu, K. and Yoshikawa, M.
Identification and characterization of virK, a virulence-associated
large plasmid gene essential for intercellular spreading of
Shigella flexneri
JOURNAL Mol. Microbiol. 6 (16), 2387-2395 (1992)
MEDLINE 93023866
PUBMED 1406277
COMMENT Submitted (27-APR-1992) to DDBJ by:
Noboru Nakata
Institute of Medical Science
University of Tokyo
4-6-1 Shirokanedai, Minato-ku
Tokyo 108
Japan
Phone: 03-3443-8111x252
Fax: 03-3443-3693.
Location/Qualifiers
1. .1642

FEATURES
source

/organism="Shigella flexneri"
/mol_type="genomic DNA"
/db_xref="taxon:623"
/plasmid="Plasmid pMYSH6000"
373. .1323
/gene="virK"
373. .1323
/gene="virK"
/codon_start=1
/transl_table=11
/product="virulence protein"
/protein_id="BAA01778.1"
/db_xref="GI:216944"
/translations="MFSVSNLSPIGFKRIVFSSDSLPGKWEHRKFRFVYLRCALNP
VASIRYVYELRSIQCEIILAIQPTLPARLHRYLKHGGRMSRGYILHEVRFVONL
PEKYSFELPQKSVSLVQFIGNKDGEDFDIQCSFGDFRGELMSLFFPKVILARLTF
SVLTQNGHTAFITGGQAPKNTGPDIIICATRCACYGKFLPKRIIFPAFCALMKACNV
ECLAVSEHSVFRQLRYQKRITFVAVYDFWESVAGKTCGDMYKLPQVVRKPLSN
IASKKRSEYRKRYALLDYIHTAIRSLDAYPVNSEHQDLN"

ORIGIN

Query Match 13.5%; Score 159.4; DB 1; Length 1642;
Best Local Similarity 65.1%; Pred. No. 6.5e-40;
Matches 235; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 210 TGACTAGCCAGACACAAATGATCTGTGCGGTTCTGTTAATATCAAAACCGGTACTCAATA 269
Db 361 TGGTTGCTTAATGTGAGTATTTTAAACGAGTTTTTATTAATATCAAAACCGCTCTTCTATA 302

QY 270 TCTTCTTGGCGTGGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGCGC 329
Db 301 TCCTGTTTTGCTGGCGAGCCATCTGTATAAAACGCCAGGGTCATTAAAAAAATCATTC 242

QY 330 AGTGGCGCGGTCCATGCAGACACATCCCCACCGGTAACAGCGTCCCTGTCACTTCTTC 389
Db 241 AGGCGGCACATCCATCGGTCTTATTACCTGCTGGTAAACAATGTCCCGGTCTGGTTATTTC 182

QY 390 TGAATGACATCAGGGATCCCGCGCTCTCACTGGCGATAACCGGCGACCGCGGAGACTGCAC 449
Db 181 TGAATTAACATCAGGATTCACCAATCTGACTGGCCAGTACAGGCACAGAAATGCGCAT 122

QY 450 GCTTCAGCCAGTACCAATACCAACCGTTTCATTTTCCGAAGGCATGACCAACACACTGGCA 509
Db 121 GCTTCTGCCAGTACCAATACCAACCGATTCTGTTTCTGAAGGCAGAACCAACAGACTGGCA 62

QY 510 ATCCGCTAGACCGGTAAACGCTGGGAAAGGGCACCTGCCATTAAACATCTCCGCTCATT 569
Db 61 ACCCGATATACGGGGGCGGAGAAAAAATTTGTACGAATAAAAAACATCATCATCATG 2

QY 570 C 570
Db 1 C 1

RESULT 16

E06635/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

1 (bases 1 to 1642)
Yoshikawa, M.
ATTENUATED LIVE VACCINE OF DYSENTERY BACILLUS HAVING HIGH SAFETY
Patent: JP 1994030766-A 1 08-FEB-1994;
RES DEV CORP OF JAPAN
OS Shigella flexneri 2a
PN JP 1994030766-A/1
PD 08-FEB-1994
PF 15-JUL-1992 JP 1992210772
PI YOSHIKAWA MASANOSUKE
PC C12N1/21,A61K39/00,A61K39/00,A61K39/112,A61K39/112,C12N15/31,
PC C12P21/02,
PC (C12N1/21,C12R1:01), (C12P21/02,C12R1:01);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: strain=YSH6000T;
CC *source: clone=pMYSH6000;
FH Key Location/Qualifiers

FT mat_peptide 373. .1320
/product="Shigella virK protein".

FEATURES
source

1. .1642
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 13.5%; Score 159.4; DB 6; Length 1642;
Best Local Similarity 65.1%; Pred. No. 6.5e-40;
Matches 235; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 210 TGACTAGCCAGACACAAATGATCTGTGCGGTTCTGTTAATATCAAAACCGGTACTCAATA 269
Db 361 TGGTTGCTTAATGTGAGTATTTTAAACGAGTTTTTATTAATATCAAAACCGCTCTTCTATA 302

QY 270 TCTTCTTGGCGTGGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGCGC 329
Db 301 TCCTGTTTTGCTGGCGAGCCATCTGTATAAAACGCCAGGGTCATTAAAAAAATCATTC 242

QY 330 AGTGGCGCGGTCCATGCAGACACATCCCCACCGGTAACAGCGTCCCTGTCACTTCTTC 389
Db 241 AGGCGGCACATCCATCGGTCTTATTACCTGCTGGTAAACAATGTCCCGGTCTGGTTATTTC 182

QY 390 TGAATGACATCAGGGATCCCGCGCTCTCACTGGCGATAACCGGCGACCGCGGAGACTGCAC 449
Db 181 TGAATTAACATCAGGATTCACCAATCTGACTGGCCAGTACAGGCACAGAAATGCGCAT 122

QY 450 GCTTCAGCCAGTACCAATACCAACCGTTTCATTTTCCGAAGGCATGACCAACACACTGGCA 509
Db 121 GCTTCTGCCAGTACCAATACCAACCGATTCTGTTTCTGAAGGCAGAACCAACAGACTGGCA 62

QY 510 ATCCGCTAGACCGGTAAACGCTGGGAAAGGGCACCTGCCATTAAACATCTCCGCTCATT 569
Db 61 ACCCGATATACGGGGGCGGAGAAAAAATTTGTACGAATAAAAAACATCATCATCATG 2

QY 570 C 570
Db 1 C 1

RESULT 17

AF326777 AF326777 66714 bp DNA linear BCT 18-JUN-2002
LOCUS Shigella flexneri 2a SRL pathogenicity island, complete sequence.
DEFINITION AF326777
ACCESSION AF326777
VERSION AF326777.2 GI:21450881
KEYWORDS
SOURCE Shigella flexneri 2a
ORGANISM Shigella flexneri 2a
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Shigella.
REFERENCE 1 (bases 1 to 66714)
AUTHORS Luck, S.N., Turner, S.A., Rajakumar, K., Sakellaris, H. and Adler, B.
TITLE Ferric citrate transport system (Fec) of Shigella flexneri 2a
YSH6000 is encoded on a novel pathogenicity island carrying
multiple antibiotic resistance genes
JOURNAL Infect. Immun. 69 (10), 6012-6021 (2001)
MEDLINE 21437601
PUBMED 11553538
REFERENCE 2 (bases 1 to 66714)
AUTHORS Luck, S.N., Turner, S.A. and Rajakumar, K.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2000) Microbiology, Monash University, VIC 3800,
Australia
REFERENCE 3 (bases 1 to 66714)
AUTHORS Luck, S.N., Turner, S.A. and Rajakumar, K.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2002) Microbiology, Monash University, VIC 3800,
Australia
REMARK Amino acid sequence updated by submitter
REFERENCE 4 (bases 1 to 66714)
AUTHORS Luck, S.N., Turner, S.A. and Rajakumar, K.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-2002) Microbiology, Monash University, VIC 3800,
Australia
REMARK Sequence update by submitter
COMMENT On Jun 18, 2002 this sequence version replaced gi:15808696.
FEATURES
source
1. .66714
/organism="Shigella flexneri 2a"
/mol_type="genomic DNA"
/db_xref="taxon:42897"
/note="corresponds to Escherichia coli min 23"
misc_feature 458. .66714
/note="SRL pathogenicity island"
gene 619. .1836
/gene="int"
CDS 619. .1836
/gene="int"
/note="member of the P4 bacteriophage integrase family"
/codon_start=1
/transl_table=11
/product="integrase"
/protein_id="AAL08426.1"
/db_xref="GI:15808697"
/translation="MAVLTDTKARIKDDKPLPHGGITGLTLHPSSVKRGKVFYR
VSPVTKRRNAGLTYPEVSTAEARTARIMRQLAAGDDPLEIKKAESKVIPTFA
DAARVRAHSPGENPKHVRWMLSTLENTAFQGLAKTLDSTADVAETLRPVLMT
LSEATSVKORIHVMQWAGHFCVANPDVVDHLLPQOTGRGRDEHPAMPWRQLPL
FVATSVYDPYVNTVTRALLMVLITATRSGEARGMRAEIDFHKRVWTI PAERKARL
QHVPLSRQALYILENIRGLHDELVPSPKQQLSDMLVLTSLRKKKAVSDIPGRVA
TAHGPRSTFRDWCSEQSYSDRLAERALLAHTLKKKVEAYHRTDLLEQRPVPMQAWADY
VMSQLAKNTEFA"
complement (2006. .3535)
CDS /note="orf2; similar to yfj1 on the cryptic P4 prophage
CP4-57 of Escherichia coli K-12"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAL08427.1"
/db_xref="GI:15808698"
/translation="MMTTLSPLELDTTKLPDRCOALLDSEHSETGIRREILLVMTL
VMAASVQDTHEVELSGQRTSLQIFMCLSSASGGKTSACAKLIAPVHETERELHONY

IDKQVDRMMEMWTTDKLIRRYKKEIERSPENAAAAAALKEKCIANKPEPPVQV
LIYNDATPEGIALKLSQSPSLLLSDEGGTILDKRPERKALYNTLWSGQFVNVHRS
RPFRLNDVRLTILITQPVIFNKFPTLTGDTGRNGFLARLIFCEGAIKIMTTEPH
TAPVVTQOCASCFRKSGQISRLSRASRRARRAKERICMTLSNAASRALERPHK
ENMDTVRONPHMGTFEDIIVRKEOVVRIIAALLEKDPDSTVITLSTNSAILYIEF
YFKHLIYKLESLEISPAEKLDKWLQKRIITAGYIPQKSYILOQYAPALRKKCVLDE
ALDLIAQRKIRIDNDLVYIGNTITPSELAKLNIIPAFDAGVFIQCHQNLKLRNR
L"
complement (3507. .3812)
/notes="orf3; similar to AlpA on the cryptic P4 prophage
CP4-57 of Escherichia coli K-12"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator"
/protein_id="AAL08428.1"
/db_xref="GI:15808699"
/translation="MEMNMNSMOTGOLYTPSRILRMAELTRLIGISRSTIYVKINPA
SKYHDPFPKPIRLGSASVGRASADLWMLHTAPAWSEPNKLELNDDDKTITP"
3868. .5730
/note="orf4"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAL08429.1"
/db_xref="GI:15808700"
/translation="MQTNRRAQHMRTFTNNLYDICTVLGLFKEGENPAHKRKTNPFM
HOKFWQRYNEISRIIDAEQVFSQORRIIYARVEHYFYMNNSYVHSTILKPELRSY
CLRTFGVILVDMYNTYRPENDSAFYHILNFLOKSYCPCLDHADTESBAAVRVL
REYLAELGENKEDPHENKLYALGKYGTIRKDKGSKSLMOQYIMAKINEYKVDRE
KLLDDELEKVLNIDKFYNAFYSLVLLDIQRTKILQSLAYILRVLVREGLMIHGL
YGAAYLYDFTSFTTPYAKLLEMFYSLVSLDLSRYSVSLDDKSQBYISRLKQ
LVFNINDKNGCDDAYLKKIISYFQQLNEAVHTVSCYETLAVYICLRKKNINDLOH
YDDMERKGLFGELPSGVYRGALSLRLTALEKVNKNKIYKGSFLFYLDHYKAYQDAFI
EKIPLIDPVYKEGIOVDANNFTLMRVIKMNCWLEKISTKPIAVPYITGLLDDVEK
VLQINILDKVEYVDGKTLAEVIMENKVLSSRKETMIGLFTGSKKYTLQOCVEKL
GULVHYVSPVDEIKVMMLYGDKAENRRNRMYYDLTIICEDDIRNPPELS"
7104. .7697
/notes="orf5"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAL08430.1"
/db_xref="GI:15808701"
/translation="MTNANPNYKMDFMILKDVNHEMDGMFQRFSLKLPFRIDFAYRKD
TPSFGHSCKSHMCEIYRLISETQTLAGYVYVMEYTPDKGLHITHFTGYLDGQRHKS
YQISROLGDIWREITTEGEGVHLCRAKDKYPIVDHVIHYSDKSASVDDLYALSYLAK
QDQKHGIIIGRSRLPEKSNRGRPDITELMGLMSELS"
8449. .8745
/notes="orf6; similar to yfj3 on the cryptic P4 prophage
CP4-57 of Escherichia coli K-12"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="unknown"
/protein_id="AAL08431.1"
/db_xref="GI:15808702"
/translation="MSLSRLPPFNKDAYVHLGNYSRSDTLURGLITGAWYSALGLELED
HQGLHFPFANGRYVILINDPNPEQYQRLVRLDYLTKLDTKASGQRNFYSGF"
complement (9112. .10421)
/notes="orf7; similar to a LysR-like transcriptional
regulator"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator"
/protein_id="AAL08432.2"
/db_xref="GI:21450882"
/translation="MKNIEYKWLVDFTLEACHFHSQAERNLSQPAFSRRIKALEA

	CDS	AIGVVLFDRTTTPLOLTERGKLFPHSQTRSLQLQOECNLGELNGQSLLGVPNIKTAAXH SUSLSVPKLXLTAYGEFVVHYEADIVQAVNTAREGSKDFIISFRDELDMGSPF CLKGFESLYPVCADAGHPVFDTIQPVFLNATYTSYNGLRVNRHLAEVGQTG RITFSSELKNWALNGYGAWLPINWSIVDELQTKILCDAAKLVPIQAIIYRM NRLNSTASLNWRILLQEHPMDDLIOQISMEEPARN"
		12299..12988 /note="Orf8" /codon_start=1 /evidence=not_experimental /transl_table=11 /product="unknown" /protein_id="AAL08433.1" /db_xref="GI:15808704" /translation="MGPANTVDLENKFVTFTAACDQEHIPLIISSIDPIDRDTDALM HGRSPLFWRDYMHKLEDGAEICVIPCNTAHYFKELDKACHIDILISIVETTINEV RACKGTIGLLATNYLMGYLKIESGIFCVSPDADQKKVMESIYSUKAGNIHA AOKMNEQAEIILFSRGAEIILVLGTEIPIVILAQAVEQPSPRYIDSTASLVRAGIKWE NRVKHHYLITQ"
	CDS	13054..14352 /note="Orf9; similar to DcuA in Escherichia coli K-12" /codon_start=1 /transl_table=11 /product="putative anaerobic decarboxylate transporter" /protein_id="AAL08434.1" /db_xref="GI:15808705" /translation="MVWLBIIVLGAIPIFGIRGGIGIGLCGLGLAILTLFGFLPMG SPVDVILIMTVVAASAQAAGGMVLYRVAGNMERNPKYINIAPITWSMTIN AGTFPIVSFTLVIAEVAKESGRPSRTLGSVASOVAISGPSISAMAAMLTIMEV NGSFIVQMSSCLPTSFAANMAAFIASRQEGQODEVIELRPGLKGIVKTENNSSI KEGATSUGFLLLATAIVILAAPQLRPGFDIGMETRIDIIIICMLSAACLMMVLCL KTADAIIITSFRAGMSLAVALIGLVIITGLTDIIDLHTLETDIADGDILOAYPMLLAL VLFGTTMCVLSOGATTPLIPLIALADVPTWAIIASYYAVTGFFVLTYPTTSLAMEF DTTGTTRVCKYVLNHPPILGGLIAGVALGFVIAPMII"
	repeat_region	complement(14553..>15322) /note="remnant of IS600-like element; central region deleted"
	repeat_region	complement(15323..16090) /note="IS1 element"
	gene	complement(16171..16962) /gene="aadAl"
	CDS	complement(16171..16962) /gene="aadAl" /note="confers streptomycin and spectinomycin resistance; similar to AadAl from Tn21 on plasmid R100" /codon_start=1 /transl_table=11 /product="streptomycin adenyltransferase Aadal" /protein_id="AAL08435.1" /db_xref="GI:15808706" /translation="MRVIVAESVTSQISEVGVIERHLEPTLAVHLVGSADVGLKP HSDIDLVTVLRBDETLAGRIINDDLETSSAPSESEILRAVEVTIVVHDIIIPWRY AKREILOFGSWENNDILAGIFEPAITIDIALLTKAREHSIVALVGPAEEILFPVPEQ DLFXELNETLWNSSPDWKDXRNWXTLSKIWXSVTXTKIAPKDVAADMWERLPA QXPQVILEARQXYLGQEDERLASRADQLEEFEHVYKGETIKTVVGK"
	gene	complement(17075..17905)
	Query Match	13.1%; Score 154.8; DB 1; Length 66714;
	Best Local Similarity	58.2%; Pred. No. 4.7e-38;
	Matches	292; Conservative 0; Mismatches 207; Indels 3; Gaps 1;
Qy	683	CATAAATCATGCCTTTTTTCGGGACGTAGCACTCCACCCTGAACGATGAAGCGAACATT 742
Dd	31985	CCTAAATTCGTGCCCTTTTTTCGGGGCGCAACTCCTCTGAGCAAATTACGGGGCATCCC 32044
Qy	743	GTCGTCTGATGCAGCCAGCGGTGGAATATGCAGGGGTAAACGGTCGCGTTCGTTCAATTATG 802
Dd	32045	TCTTCGGGAGCCGAGCCAGGAGAACAACATCGGAGSGCAAACTGTTTCGGGAATCAACGTA 32104
Qy	803	CAATCGGGCGAGTCGAAAACCGCGTAGAATAACCGTTACCGGTGTCTGCACACCTTCGCG 862
Dd	32105	TAAATTCCTGGAATCAAGCCAGGAGGCCAACAGTAATCCCCAGATTGCGACATCTCTTCGT 32164
Qy	863	CATCACATCGCCATCATCGGGTGAGATAGSCAACAAATGAATCACACAGATAAATTCAG 922

Db	32165	CTCCAGATGTCCTTCATATTTGTTCCGGGAACAATCACTCATCGAAAATGATTTAT	32224
Qy	923	GGAAAAAGCTTCTGTCCTTACGGGTGATGTAGGTTTTTTTGTCTGCAAAATAGTGAAGCGGTG	982
Db	32225	TGAAAAAATTTTGTGTTTGTGTCGTAATATGTCCTTTTGGCCGTGATAATTCCTGTAATGGATG	32284
Qy	983	----ACAGCATATCAGACGGCTCAGTCCGTCTATATTAATCTGTCATGGCCACTATGGCAGAT	1039
Db	32285	TTTTCAGGTGTAATAAACCGTACCAACCAACAATATTAATCTATCATGTCACATGTCACAAAC	32344
Qy	1040	GACCATCATCAGGTTTAAATTCCTCCGATATTCGTCGAAAGTCTGAGATGGAAGGAAGGTG	1099
Db	32345	GATTGCAATCGGCTGAAACACTATGAATTTCCGAGTAACTCTCCATACAGTAGGATGTG	32404
Qy	1100	AAGGCTGTTCCTGAAAGGAATAAAGTGAACATCATGCCCTCTTTTCTGCGTCTCCGGAGC	1159
Db	32405	AAGACTGTTTTCGAAATAACGCGAATGTGATATCAATTCCAAACTTACTTGGCTTCAAAAGC	32464
Qy	1160	AAATTTTACTTTTTTCTCTGCAG	1181
Db	32465	AAATTTGCTGTTCTCCCTGCAG	32486
RESULT 18			
AE016912/c			
LOCUS		300774 bp	DNA linear BCT 14-OCT-2003
DEFINITION		Chromobacterium violaceum ATCC 12472 section 3 of 16 of the complete genome.	
ACCESSION		AE016912	AE016825
VERSION		AE016912.1	GI:34101873
KEYWORDS			
SOURCE		Chromobacterium violaceum ATCC 12472	
ORGANISM		Chromobacterium violaceum ATCC 12472	
REFERENCE		1	Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales; Neisseriaceae; Chromobacterium.
AUTHORS		1	(bases 1 to 300774)
CONSTRM			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
1			Brazilian National Genome Project Consortium
2			The complete genome sequence of Chromobacterium violaceum reveals remarkable and exploitable bacterial adaptability
3			Proc. Natl. Acad. Sci. U.S.A. 100 (20), 11660-11665 (2003)
4			22882880
5			14500782
6			2 (bases 1 to 300774)
7			Vasconcelos, A.T.R., de Almeida, D.F., Almeida, F.C., de Almeida, L.G.P., de Almeida, R., Gonçalves, J.A.A., Andrade, E.M., Antonio, R.V., Araripe, J., de Araujo, M.F.F., Filho, S.A., Azevedo, V., Batista, A.J., Bataus, L.A.M., Batista, J.S., Belo, A., vander Berg, C., Blamey, J., Bogo, M., Bonato, S., Bordignon, J., Brito, C.A., Brocchi, N., Burity, H.A., Camargo, A.A., Cardoso, D.D.P., Carneiro, N.P., Carraro, D.M., Carvalho, C.M.B., Cascardo, J.C.M., Cavada, B.S., Chueire, L.M.O., Pasa, T.B.C., Duran, N., Fagundes, N., Falcao, C.L., Fantinatti, F., Farias, I.P., Felipe, M.S.S., Ferrari, L.P., Ferro, J.A., Ferro, M.I.T., Franco, G.R., Freitas, N.S.A., Furlan, L.R., Gazzinelli, R.T., Gomes, E.A., Gonçalves, P.R., Grangeli, R., Grattapaglia, D., Grisard, E.C., Guimarães, C.T., Hanna, E.S., Hungria, M., Jardim, S.N., Laurino, J., Leoi, L.C.T., Passarella, L., Lima, A., Loureiro, M.F., Lyra, M.C.P., Macedo, M., Madeira, H.M.F., Manfio, G.P., Maranhão, A.Q., Martins, W.S., di Mauro, S.M.Z., de Medeiros, S.R.B., Meisner, R.V., Menck, C.F.N., Moreira, M.A.M., Nascimento, P.P., Nicolas, M.F., Oliveira, J.G., Oliveira, S.C., Paixao, R.F.C., Parente, J.A., Pedrosa, F.O., Pena, S.J.D., Pereira, J.O., Ferreira, M., Pinto, L.S.R.C., Pinto, L.S., Porto, J.I.R., Potrich, D.P., Neto, C.E.R., Reis, A.M.M., Rigo, L.U., Rondinelli, E., dos Santos, E.B.P., Santos, F.R., Schneider, M.P.C., Seanez, H.N., Silva, A.M.R., da Silva, A.L.C., Silva, D.W., Silva, R., Simoes, I.C., Simon, D., Soares, C.M.A., Soares, R.B.A., Souza, E.M., Souza, K.R.L., Souza, R.C., Steffens, M.B.R., Steindler, M., Teixeira, S.R., Urmenyi, T., Vettore, A., Wassem, R., Zaha, A. and Simpson, A.J.G.
8			Direct Submission
9			Submitted (22-JAN-2003) Labinfo, LNCC - Laboratório Nacional de
10			JOURNAL

FEATURES source	Computacao Cientifica, Rua Getulio Vargas 333, Petropolis, RJ 25651070, Brazil	CDS	complement (3594..4337) /locus_tag="CV0569" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AA058245.1" /db_xref="GI:34101878" /translation="MKSFRKLGRALAFIDPIVVLSEILVVAIASNLTSYKVIQVFG TEASLGTFTMPIILMLNLAIEVYKGRSNOILPATCLAILFAVSTALUSQSDRCG LLPABRGHCEALNQSVLISEHIVRGSIFAICGLIGSQLNTRFLYLLKLSMSRLY FIRDFSSVIGEVVYTAICFMIAFYGVFPATILKIFAFSLMFKFSSTVILSWISQVI VLLIYRYOAWSEGAQGRKVKFSSRYLQV" complement (4334..4861) /locus_tag="CV0570" complement (4334..4861) /locus_tag="CV0570" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AA058246.1" /db_xref="GI:34101879" /translation="MQAKNLLKPHSVLHEFKLNALRRIRDKLISISLARITGIKQST LHRGLFEDRELTFESNAHAIGHALGINLGGGETPRMAPIVTGLDOLDALSGRQPVWDE FILLPSLDLSALAVDQSLFRQRLFPRAVVVLQTAHPRGRRLVYPDHGALSLEDSGR SALGAVAGILFRKNA" 4993..5178 /locus_tag="CV0571" 4993..5178 /locus_tag="CV0571" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AA058247.1" /db_xref="GI:34101880" /translation="MIYVICGMLVWLKEINPDSLEGGFQSGSLVSHDYFCEMSCLK LSGYLRPLRVFLILCF" 5297..5863 /locus_tag="CV0572" 5297..5863 /locus_tag="CV0572" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AA058248.1" /db_xref="GI:34101881" /translation="MRSAFVQTASGRYLDVQDPQDDIDIADIAHLAHLRCRFGGACR QYVSEAHAVRMAEILPRLKLAGLYGAKA VAGAVRPGQAGPAPQRLTAOVA RGLGHISGDDRAELAMADARLLATERDLMPDNIEWPILTGVQPLAERIWPMTIAQA LASFCAYQCRAWSSSAQIRLRQPSVR" complement (5930..7795) /genes="kup2" complement (5930..7795) /locus_tag="CV0573" complement (5930..7795) /genes="kup2" /locus_tag="CV0573" /notes="identified by sequence similarity; putative; ORF located using Glimmer/Genemark/Blastx/COG3158/TC:2.A.72.1.1" /codon_start=1 /transl_table=11 /product="potassium uptake protein" /protein_id="AA058249.1" /db_xref="GI:34101882"
	1..300774 /organism="Chromobacterium violaceum ATCC 12472" /mol_type="genomic DNA" /strain="ATCC 12472" /db_xref="taxon:243365" complement (68..598) /locus_tag="CV0565" complement (68..598) /locus_tag="CV0565" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAQ58241.1" /db_xref="GI:34101874" /translation="MLTDLTLGKAGAAPSPEQPLEMLQACHKIRRFCDOLDKLPYVI BEHGVDAARNTIDVRYPDVAGPAHHTDEEELFPLIDERVEPTASSRLEQLSAEHG YLHSCWNAIRDLLVALRDGRKIQISRNELQFARQYREHAAAEANLPTAASALSPE ELRQAGANMARRRODA" 1267..1812 /genes="ampD" /locus_tag="CV0566" 1267..1812 /genes="ampD" /locus_tag="CV0566" /notes="identified by sequence similarity; putative; ORF located using Blastx/COG3023" /codon_start=1 /transl_table=11 /product="anhydro-N-acetylmuramyl-tripeptide amidase (ampD protein)" /protein_id="AAQ58242.2" /db_xref="GI:34330281" /translation="MTLGADGWQARQLSPNCDSRPGCAPELLVIHNSLPPVRY GGAGVEQLFNSQLPDRHPYKGIQLRVSSHFIIRDGQLQFPVPGKAWHAGVSN WRGKGNPDSIGVEMEGCDFEPPFSEAQYRTLAALSRELRAULPSAIAGHEIAPGR KTDPGMFDWRRAQADSGLGF" complement (1817..2272) /locus_tag="CV0567" complement (1817..2272) /locus_tag="CV0567" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AAQ58243.1" /db_xref="GI:34101876" /translation="MSALTIEDIDADAITIRLVNTAKTYLVPLDRVIRLQCLQGLPRD EILLAGVAGKVARRRPALLAQDFSDVAVCGDFRQPHVVFDRSVGFVAGGETRVDLP FRELLGNLDAISRLSDEAAFYLGFSYALFALPGEUSFGDHPRPYLLK" complement (2269..3597) /locus_tag="CV0568" complement (2269..3597) /locus_tag="CV0568" /EC_number="4.1.3.27" /notes="identified by sequence similarity; putative; ORF located using Glimmer/Genemark/Blastx/COG0147" /codon_start=1 /transl_table=11 /product="probable anthranilate synthase" /protein_id="AAQ58244.1" /db_xref="GI:34101877" /translation="MKNRILHVSVPAPADGRQTPLFALRCNVDYSGKYALHERAVERAA LSALFFADRLALLPRSGNVRVIAETLAPLLGLSPDPAAPGAFRLPLDPAFLAERLQO LLDLPAREFDAPRGLGPGFCYELLHAKLGNAGLLGVHFVPERLLVDGAVDCLDLP ADAGGSEWPLALGPMIGADDPAAGRGLPYERMYAASRRR1ADGNLASVNSPAIQRPC ADARYR1HQDLARNPAPNLYFDGGDFQLAGSPAMFLRGRRLATSPICTGIARG RGEADEAQVATLASDKDFEECEVRADMQAKESCEDIRVDIEREVERFANVFT SASVSRLEKFGCTLADACDHLWPATVVGTPVDAAARLLAEHSRRWYAGAFGLCAN EVELGTIRITALADGMATRVGSTLSARSSPQLEQSEAKASLLILGLV" complement (3594..4337) /locus_tag="CV0569"		


```

VNHRYGKPKKATTEDDKSMEEIGDVLVFLVCLANSLLDISLEEAHDRVHMKFNTROKD
RWRKRGK"
1358. .2101
/gene="dapB"
1358. .2101
/gene="dapB"
/functions="lysine biosynthesis"
/notes="28.2% of identical aminoacids with the Escherichia
coli dihydrodipicolinate reductase; putative"
/codon_start=1
/transl_table=11
/product="dihydrodipicolinate reductase"
/protein_id="AAA92874.1"
/db_xref="GI:755603"
/translations="MSNETIKLVIAGRGRMGQEAVKLAERTPHFDLVGADHTYQQ
KLSDMVPESDAFYTDIHAFCFTQPDVLIDLTTPGKVTIKALEHGVRRVPGTT
GFSADMLKELTSTEEKGIGAIITAPNPAIGALMFKSMAANYFDEVEIIEHLHQK
LDPSGVALTAEMIISVRKEQKHPEKEILLPGARGAQNGIRLHRLQIRHDSYNR
ASPMGKLSVEQWMDQLVYGLEIID"
2117. .2530
/gene="joJf"
2117. .2530
/gene="joJf"
/notes="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92875.1"
/db_xref="GI:755604"
/translations="MKTALIAHDKKQDMQVFTTAYRDIILKNHDLVATGTTGLKIHFA
TGLQIERFQSGPGDQIGQIGALIAANALDLVIFLRDPLTAQHPDVSALIRLCDVYS
IPLATNMGTAELVRLDEGVFFERDLRGEENPV"
2523. .3197
/gene="joJG"
2523. .3197
/gene="joJG"
/notes="putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92876.1"
/db_xref="GI:755605"
/translations="MYNADVLAFGMSDDVEIGMGCTIAKFVKQEKVIMCDLTEL
SSNGTVSLRKEBAEAAIRILGADKRIQTLTPDRGLMSDQARSIVTVIRICRPKAVF
MPYKDRHPDHGNAALVERBAIFSAHKKYDEKSLPAHKVSKVYMYMINGFHQDFV
IDSIDTTEKIGSLNAYKSQFIPSKDSVSTPLTNGYIEIIVEAREKLYGKAGVEYAEV
SPFNGC"
3229. .4302
/gene="joJH"
3229. .4302
/gene="joJH"
/notes="21.6% of identity to trans-acting transcription
factor of Saccharomyces cerevisiae; 25% identity to sucrose
synthase of Zea mays; putative"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAA92877.1"
/db_xref="GI:755606"
/translations="MRKLGKLAEGHEIHFTSSIPFRLTYHPNIHPEVEVNOYA
VFKYPPDVLILAKIAEVAERENLDIIHAHYALPHAVCAYLAKQMLKRNIGIVTTLHG
TDITVGLPSLDLIRFAIESSDRVTVASSALAAETDUIKPEKTIETIYNFIDERV
YLNKNTAAIKERKIGLDEKVIHVSFNRKRVQDVIRVFRNIGTKAKLLVGDG
PEKSTACELARKTGLQVLMGNQDRVEDLYSIDLKLKLSSEKSFGLVLEAMACG
VPCIGTNGIGIPEVINKNVSGFLVDVGDVTAATARAMSILEDEQLSNRPTKAAIEMLE
NEFSKKIVSQEIQYADLAEP"
4307. .5440
/gene="paps"
4307. .5440
/gene="paps"
/notes="38.9% of identical aminoacids with the Escherichia
coli pcnB gene product, encoding poly(A) polymerase;
putative"

```

```

/codon_start=1
/transl_table=11
/product="polyA polymerase"
/protein_id="AAA92878.1"
/db_xref="GI:755607"
/translations="MEKVFIKALPVLRIIEAGHOAFVGVGAVRDSYMKRTIGDVIDIA
TDAAPQOERLFORTVDVCKDESVDYDFRPSVQVFISSLEEDLKRPDLTINAMAMTA
DGKVLDFGKKDIDQKIVRTVGKPEDRFQEDALMLRAVRPMSQLGFTLSPETEEAI
AKESKLSLHVSVERTKIEFEKLQGRASQALQTLITRILYBELPGFYHRENLISYS
BFPFSLTSREELWAAALLNLGIVLKDAPLPGKAWKLPKVIKEAIIHADIATFQGSLLDA
MTWRAGKCALLSAAKISQLRQNEKLDEKLDKIQAYQNLPIKSLDKLDITGKDLIA
LRNRPAKWVSVBELQWIEQAVVTGKLSNQKKHIEEMLKTCGQH"
5425. .6342
/gene="birA"
5425. .6342
/gene="birA"
/functions="bifunctional biotin operon repressor"
/notes="26.7% identity to the Escherichia coli bifunctional
biotin operon repressor"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="biotin [acetyl-CoA carboxylase] ligase"
/protein_id="AAA92879.1"
/db_xref="GI:755608"
/translations="MRSTLRKOLIELFQSGNEFISQKISDALGCSRTAVWKHIEL
RKEGYEVAERRRGYRLIKKPGKLSSEIRFGLKTEVMGQHLIYHVDLSSTQTAHEL
ANNAPEGLTEGVWMSLIRPDIPLQKTPQLTLLAAVAVQGIIEAAGIOTDIKWP
NDILINGKKTGILTEMSQABEDRVRSVIGIGINVQNPDPDELKDIATLSQAAG
EKIDRAGVIOHILLCFEKRYRDYMHGFTPIKLWESYALGIGTNMRARTLNGTFYCK
ALGIDDEGVLLLETNEGIKKIYSADIELG"
6501. .6533
/terminator
/notes="putative"
ORIGIN
Query Match 5.4%; Score 64.2; DB 1; Length 6540;
Best Local Similarity 54.4%; Pred. No. 1.4e-08;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 335 GCGGTCCTCATGCAGACATCCCCACGGGTAAACAGCGTCCCTGTCTCATCTTCTGAAT 394
Db 4158 GGCTGTTGGCGGCTAAACATCACCGACATCCACAAAATCCGCTCACATGTTTTTAT 4099
Qy 395 GACATCAGGATCCCGCGCTCTCAGTGGCGATAACGGGCAACGGCGAGACTGACGCTTC 454
Db 4098 AACCTCAGGATACCCGCAATGTTGTTCCAATACAAGGCACTCCGCAAGCCATCGCTTC 4039
Qy 455 AGCCAGTACCATACCAAGCGTTTCATTTCCGAAGCATCACACCACTGGCAATCCG 514
Db 4038 AAGCAGACAGGCCAAGCTTCTTTTCAGATAGCAGCGCTTCAATCGCTAATAGA 3979
Qy 515 GTAGACCGGTAAACGCTGGGAAAAGGCGACCTGCCATTAAACACATCTCCGCTCATTC 571
Db 3978 ATAAAGATCTTCAACACGGCTTGTGATTTCCAAGCATTAAGACTTGGTCTTCCAGCC 3922
RESULT 20
AX100832
LOCUS AX100832 10801 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 76 from Patent WO0121772.
ACCESSION AX100832
VERSION AX100832.1 GI:13619763
KEYWORDS .
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Yocum,R., Patterson,T.A., Hermann,T. and Pero,J.G.
TITLE Methods and microorganisms for production of panto-compounds
JOURNAL Patent: WO 0121772-A 76 29-MAR-2001;
OMNIGENE BIOPRODUCTS, INC. (US)
FEATURES Location/Qualifiers
source 1. .10801

```

```

/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Recombinant pAN240 plasmid"

ORIGIN
Query Match      5.4%; Score 64.2; DB 6; Length 10801;
Best Local Similarity 54.4%; Pred. No. 1.5e-08;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 335 GCCGGTCCATGCAGACATCCCCACCGGTACAGCGTCCTGTGCATTCCTTCGAAT 394
Db 6490 GGTGTTGCCGCGGTAACTACCGACATCCACCAAAAATCCGCTCACATGTTTTTAT 6549

QY 395 GACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGCAGCGACGCCGAGACTGACGCTTC 454
Db 6550 AACTCAGGATACCGCAATGTTGTTCAATACAGGCACTCCGACGCAATCGCTTC 6609

QY 455 AGCCAGTACCATACCAACGCTTCATTTTCGGAAGCATGACCAACCATCTCGCAATCCG 514
Db 6610 AAGCAGGACAAGGCCAAAGCTTCTTTTTCAGATAGCAGCAGCTTCAATCGCTAATAGA 6669

QY 515 GTAGACCGGTAAAGCTGGGAAAGGCACTGCCATTAAACATCTCCGCTCATTC 571
Db 6670 ATAAGATCTTCAACACGCTTGTGATTTCGAAGCATTAAGACTTGGTCTTTCCAAGCC 6726

RESULT 21
BACYP1A/c
LOCUS
DEFINITION
BACYP1A 23775 bp DNA linear BCT 31-JAN-2003
Bacillus subtilis clone YAC15-6B ypiABF genes, qcrABC genes,
ypjABCDEFHI genes, bira gene, panBCD genes, ding gene, ypmB gene,
aspB gene, asnS gene, dnaD gene, nth gene and ypoC gene, complete
cds.
ACCESSION
L47709
VERSION
L47709.1 GI:1146223
KEYWORDS
Bacillus subtilis
SOURCE
Bacillus subtilis
ORGANISM
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE
1 (bases 1 to 23775)
Henner,D., Gollnick,P. and Moir,A.
Analysis of an 18 kilobase pair region of the Bacillus subtilis
chromosome containing the mtr and gerC operons and the aro-trp-aro
supraoperon
Proc. Int. Symp. Genet. Ind. Microorg. 6, 657-665 (1990)
2 (bases 1 to 23775)
Sorokin,A., Azevedo,V., Zumstein,E., Galleron,N., Ehrlich,S.D. and
Serror,P.
Sequence analysis of the Bacillus subtilis chromosome region
between the serA and kds loci cloned in a yeast artificial
chromosome
Microbiology 142 (Pt 8), 2005-2016 (1996)
96349105
FEATURES
8760912
Location/Qualifiers
1..23775
/organism="Bacillus subtilis"
/mol_type="genomic DNA"
/strain="Marburg"
/sub_strain="168"
/db_xref="taxon:1423"
/map="202 degrees"
/clone="YAC15-6B"
/clone_lib="chromosome YAC library of Pascale Serror"
/note="Bases 57126 to 80902 of the YAC15-6B clone"
1..33
/gene="aroE"
1..33
/gene="aroE"
/standard_name="transcription terminator"
/notes="putative"
/citation=[2]

gene
terminator

348..1619
/gene="ypjA"
348..1619
/gene="ypjA"
/function="hypothetical"
/notes="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38432.1"
/db_xref="GI:1146224"
/translation="MNTLIQEAIKLVEAGTEKGLNLTSLKAEKQLHDEKAIQAOLYY
EWGDKALISLDHLFNETLNFYAEILLIDIDSEKALAVLETIPETDPSYPE
SILLMADLYQMOGLFEVSOKLFOAKSILDNEPVIDFALGELYFAQGVAKAVQYKPT
TAEQSEIGVGNVHORLARSLSASGPFEDAIPLYKAVDENPDNTIFGVGFTALQAG
LYKTAIKQLSDLKELDPSTSLYMPLSKSYEAGMYEALKTAKEGIRIDYKNEIFL
YAAKMAKLGKSEEGKLLQALDPGFVEALHTLGLAVTHKKEEDYQIIDLIQEVR
YEEDEPKYNYLSEASAYTELEQVEEAKQFEEAALHYHREDRDFLYEASFLLEGLQKE
ALPLKKVLEMDGANEELETILRIEDEFSR"
1691..2230
/gene="ypjB"
1691..2230
/gene="ypjB"
/function="hypothetical"
/notes="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38433.1"
/db_xref="GI:1146225"
/translation="MQTPVSVNEKKDFIRWFLNHQYLRRECVMILNYLMSHDSLMK
VHFVEQAEFCPRGIMSTHCVEBVEPFRFYKENVMTDAKSPHDIRLNKQODLFIOLN
DSYKSSPBYAAVLESNPHI PKNLPENKQDQGLAEQILEHAISTFQREKLLKDDIAL
DRHDKAEFEQLSRQLNLT"
2285..2731
/gene="ypjF"
2285..2731
/gene="ypjF"
/function="hypothetical"
/notes="putative"
/citation=[2]
/codon_start=1
/transl_table=11
/protein_id="AAB38434.1"
/db_xref="GI:1146226"
/translation="MEWRITDAKDYLAQADYIDTAVIPLINIRVNNHFQMAEKGEFT
QLSEELERQLKGRVLLPPYTYVDRNETVQGLKDLREBELTEFPFHVLLTSDSWR
ASDALGKMIVTSVPLEHLNDSLKRKILDRTABEILNVLQLWSTL"
2771..2776
/citation=[2]
/evidence=experimental
2794..2799
/citation=[2]
/evidence=experimental
2874..3377
/gene="qcrA"
2874..3377
/gene="qcrA"
/function="menaquinone oxidase"
/standard_name="Rieske protein"
/notes="26.2% of identical aminoacids with the
cyanobacterium Nostoc PCC 7906 Rieske iron-sulfur protein"
/citation=[2]
/codon_start=1
/evidence=experimental
/transl_table=11
/product="Rieske iron-sulfur protein"
/protein_id="AAB38435.1"
/db_xref="GI:1146227"
/translation="MGKXHDISRROFLNYTLNGVGGFMAASMLMPVRFALDPVLKST
GKQDMVQVSVDELTKEPQRFDPKINQVDWYSESBSAWFKNGDEIVALSICKH
LGCTVWNNSDPKNPKRFPCPHGYLEKDGTVNPGTPPLAPLDHYEQEVKGLYLGK
AKPKGEG"

-35_signal
-10_signal

Gene
CDS
```


Danchin,A.
The complete genome sequence of the gram-positive bacterium
Bacillus subtilis
NATURE 390 (6657), 249-256 (1997)
98044033
9384377
2 (bases 1 to 201375)
Kunst,F., Ogasawara,N., Yoshikawa,H. and Danchin,A.
Submitted
Direct Submission
Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75724,
Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,
adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45
68 89 48
On Jul 7, 2003 this sequence version replaced gi:2634478.
This entry contains data from release R16.1 of the Subtilist
database. Further data on gene annotation and detailed information
about changes from previous releases can be found at
<http://genolist.pasteur.fr/Subtilist/>.
Location/Qualifiers
1. .201375
/organism="Bacillus subtilis subsp. subtilis str. 168"
/mol_type="genomic DNA"
/strain="168"
/db_xref="taxon:224308"
complement (246. .371)
/genes="yopL"
/locus tag="BSU20850"
complement (246. .371)
/genes="yopL"
/locus tag="BSU20850"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14003.1"
/db_xref="GI:2634505"
/db_xref="UniProt/TREMBL:O31926"
/translations="MKKLIMALVIGALGTSTYSISADSSIQQASGDYEVAGMPRGA"
complement (385. .1545)
/genes="yopK"
/locus tag="BSU20860"
complement (385. .1545)
/genes="yopK"
/locus tag="BSU20860"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14004.1"
/db_xref="GI:2634506"
/db_xref="UniProt/TREMBL:O31927"
/translations="MELIRIAMKDLNDNSLMNKWATVAGLNPNPLDYFLNHDGKT
FNEFSIVNIKSVQPDREYELMKDYLCLDVKTKAARSALYADAMFPEIEDVLID
SMISCNMSKSGKYKHIRELSNSVITEFAVKRLGKLNKTPMNSFSRLLLLYH
YLSTGNFSPAQIKQIDSEISENSMYIRNTYQTRVHVLMSNLKUNESLECREYSK
KALESNIIRFPQVSYLTGNSLFSNLAQENFLKLSISVQENYNNIQQALCF
LNNVWRKNKIINFSDTSIMDLQEAHCFINFNENSKAKEVLDKLLDLVHNDLMAH
YUUGRLQNKACFYSSIEYFKSKNDKFLIRLPLELQKNGENQKLELLLL"
complement (1722. .2138)
/genes="yopJ"
/locus tag="BSU20870"
complement (1722. .2138)
/genes="yopJ"
/locus tag="BSU20870"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14005.1"
/db_xref="GI:2634507"
/db_xref="UniProt/TREMBL:O31928"
/translations="MKNTFLYFRWEDLHGEIGVDSFNLLRASYSNLSEQOLVELIKEL
ISIERDIAAKFDIHLSENAPVDERQHVVYKGVAGDMYKMDLLSLVTALDITNTLD
HVQNIUSLAKLSRFPREIPARPAKDIAEEVYTSLK"
complement (2140. .2673)
/gene

CDS
/gene="yopI"
/locus tag="BSU20880"
complement (2140. .2673)
/gene="yopI"
/locus tag="BSU20880"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14006.1"
/db_xref="GI:2634508"
/db_xref="UniProt/TREMBL:O31929"
/translations="MNNIGELIISNFEIGIIGALLGVIVTILTHILKHFGQIKPIYVDF
BYFTKDNQDGTGNTVMPKDEAKQIEIHSQIEIYNGAEIPKVLREIKFCFYKNTNLIY
SVTPDDKATTEFAEFGYYRDKLFINLPKQIIAINIIKFLNEKETKQVKCNRYVL
EAKDHNGMKYKVLGEP"
complement (2700. .3236)
/genes="yopH"
/locus tag="BSU20890"
complement (2700. .3236)
/genes="yopH"
/locus tag="BSU20890"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14007.1"
/db_xref="GI:2634509"
/db_xref="UniProt/TREMBL:O31930"
/translations="MSDONEKSPSWGDIILKSPKYLGLAVPSGIGLGNVGLCKT
LGKDRAINSYKLYLGLVFLASTSPILSHFIWISLGKKNKIDOKYKFKOKERLRLN
RREKQILSPYIFDDVRVSVELSITDGTQAQLEHLKIIYRSSNISNRSGSYFAYNIQPWAR
GYLTNNKHLHESIDHIL"
complement (3275. .3406)
/genes="yopG"
/locus tag="BSU20900"
complement (3275. .3406)
/genes="yopG"
/locus tag="BSU20900"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14008.1"
/db_xref="GI:2634510"
/db_xref="UniProt/TREMBL:O31931"
/translations="MYWIEWIENGKKNTIVASGWIWAAILEDLYOKRFYEVWEKRL"
complement (3417. .3632)
/genes="yopF"
/locus tag="BSU20910"
complement (3417. .3632)
/genes="yopF"
/locus tag="BSU20910"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14009.1"
/db_xref="GI:2634511"
/db_xref="UniProt/TREMBL:O31932"
/translations="MPLIDYFVLOFENKEYFKAFKLDSEGYLTSSDLHEADSKIHNL
EYIEVASSELKTKCNQCEVREIQVKR"
complement (3636. .3887)
/genes="yopE"
/locus tag="BSU20920"
complement (3636. .3887)
/genes="yopE"
/locus tag="BSU20920"
/function="unknown"
/codon start=1
/transl_table=11
/protein_id="CAB14010.1"
/db_xref="GI:2634512"
/db_xref="UniProt/TREMBL:O31933"
/translations="MIGLAYFLIINLGVGLLTGIGKIFVDQVYDEEFKELMDKETAG
MERNLASUFFKKNLNVIAFFMLIGLLPLAMRITKLFKRG"
complement (4474. .4869)
/gene

```
/gene="yopD"
/locus_tag="BSU20930"
complement(4474..4869)
/gene="yopD"
/locus_tag="BSU20930"
/function="unknown"
/notes="similar to unknown proteins from B. subtilis"
/codon_start=1
/transl_table=11
/protein_id="CAB14011.1"
/db_xref="GI:2634513"
/db_xref="UniProt/TREMBL:O31934"
/translation="MSLNQKALNSYYVTLAPSSLSFGLIFGLYLFVYSGFMAIALVTI
AIIAFATIIYLFVAAPLQWLRKRRKSLINFLIYIAVAFSAFLFWFVDFPPNAL
TMRFSFYIMSIIVAFIYFWFDSIFLRN"

Query Match      5.4%; Score 64.2; DB 1; Length 201375;
Best Local Similarity 54.4%; Pred. No. 2.9e-08;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 335 GCGGGTCCATCGACACATCCCCCGGTTACAGCGTCCCTGTCTCATCTTCTGAAT 394
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148614 GGCTGTTCGGCCGTAACATCACCGACATCCACCACAAAATCCGCTCACATGTTTAT 148673

QY 395 GACATCAGGATCCCGCCGCTCTCCTGCGGATACCGGACCGCGAGACTGACGCTTC 454
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148674 AACCTCAGGATACCGGCAATGTTGTTCCAAACAGGCACTCCGCAAGCCATCGCTTC 148733

QY 455 AGCCAGTACCATACCAACGCTTCATTTCCGAAGGATGACACACACATCGGCAATCCG 514
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148734 AGCAGGACAAAGCCAAAGCTTCTTTTCAGATGACGAGCTTCAATCGCTAATAGA 148793

QY 515 GTAGACGGTAACGCTCGGAAAGGACCTGCCATTAACACATCTCCGCTCATTC 571
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148794 ATAAAGATCTTCAACACGGTCTTGATTTCGAAGCATTAGACTTGGTCTTCAAGCC 148850

RESULT 23
BX640423/c 348251 bp DNA linear BCT 06-MAY-2004
LOCUS Bordetella parapertussis strain 12822, complete genome; segment 1/14.
ACCESSION BX640423 BX470249
VERSION BX640423.1 GI:33564799
KEYWORDS complete genome.
SOURCE Bordetella parapertussis
ORGANISM Bordetella parapertussis
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.

REFERENCE
1 Parkhill,J., Sebahia,M., Preston,A., Murphy,L.D., Thomson,N.,
Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D.,
Mungall,K.L., Cerdano-Tarraga,A.M., Temple,L., James,K., Harris,B.,
Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,F.,
Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
Holroyd,S., Jagels,K., Leather,S., Moule,S., Norberczak,H.,
O'Neill,S., Ormond,D., Price,C., Rabinowitsch,E., Rutter,S.,
Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
Skellon,J., Squares,R., Squares,S., Stevens,K., Unwin,L.,
Whitehead,S., Barrall,B.G. and Maskell,D.J.
Comparative analysis of the genome sequences of Bordetella
pertussis, Bordetella parapertussis and Bordetella bronchiseptica
Unpublished
2 (bases 1 to 348251)
Sebahia,M.
Direct Submission
Submitted (06-AUG-2003) Submitted on behalf of the Pathogen
Sequencing Unit, Sanger Institute, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, E-mail: ms5@sanger.ac.uk
FEATURES
source
1..348251
/organism="Bordetella parapertussis"
```

```
/mol_type="genomic DNA"
/strain="12822"
/db_xref="taxon:519"
1..1920
/gene="gidA"
/locus_tag="BPP0001"
1..1920
/gene="gidA"
/locus_tag="BPP0001"
/notes="ortholog of Bordetella pertussis (BX470248) BPP0001
Similar to Escherichia coli glucose inhibited division
protein A GidA or B3741 SWALL:GIDA_ECOLI (SWALL:p17112)
(629 aa) fasta scores: E(): 6.8e-164, 68.52% id in 629 aa"
/codon_start=1
/transl_table=11
/product="glucose inhibited division protein A"
/protein_id="CAE39742.1"
/db_xref="GI:33564800"
/db_xref="UniProt/TREMBL:Q7W211"
/translation="MDPREFDVIVGGHAGTEALAAARAGAQTLLLTHTNITLQO
MNCPSGIGIGKHLVKGVDALGGNAIATDEAGIQFRILSSKGPVATRVQADRV
LYRNARARLEQNPWLQFQAVDDLMVQDQVVGAVTQIGLFRATVVLTAFTFLN
GLIHVGLONYSGRAGDPANSLGRLKELQPOGRLLKTGTTPRIDGRSINYSVLEQ
PGDLDPVPVFSFLGKASHQVFLPEGLNTHIYPNGVSTSLPFDQVQYELHSLPLEN
YCPSEBDKTHRPADKASHQVFLPEGLNTHIYPNGVSTSLPFDQVQYELHSLPLEN
AHLRPGYAIERYDFDPRALKSTLETKAISGLFAGGINTGTTEBAAGGLAGANA
ALQOAKBPMVPRRDDEAYLGLVDDLVTRGVTEPYRMFTSRAYEYRLSLRDNDLRLT
EGRRLLGLVDDVRWDASFRKDAVAQEVRLKSTWNPVRLPAHAAEALGKAIEREY
SLSDLRKPNVSYEALMQARTQDEGLLAGVLEDDVLABQVETQVYAGYIARQDE
VOKHLSHEOQTPADIDYDVTSLSFVRQKLTHTPETIGQAARVSGVTPAALSLLL
IHLKRLHYSRRQAA"
19..1893
/gene="gidA"
/locus_tag="BPP0001"
/notes="HMPfam hit to PF01134, Glucose inhibited division
protein A, score 0"
820..864
/gene="gidA"
/locus_tag="BPP0001"
/notes="PS01280 Glucose inhibited division protein A family
signature 1."
1917..2609
/gene="gidB"
/locus_tag="BPP0002"
1917..2609
/gene="gidB"
/locus_tag="BPP0002"
/notes="ortholog of Bordetella pertussis (BX470248) BPP0002
Similar to Pseudomonas putida glucose inhibited division
protein B GidB SWALL:GIDB_PSEPU (SWALL:P25757) (216 aa)
fasta scores: E(): 1.5e-24, 42.59% id in 216 aa, and to
Escherichia coli glucose inhibited division protein B GidB
or B3740 or Z5240 or Ecs4682 SWALL:GIDB_ECOLI
(SWALL:p17113) (207 aa) fasta scores: E(): 1.6e-12, 39% id
in 200 aa"
/codon_start=1
/transl_table=11
/product="glucose inhibited division protein B"
/protein_id="CAE39743.1"
/db_xref="GI:33564801"
/db_xref="UniProt/TREMBL:Q7W210"
/translation="MSAVPDPGGPAQRLAQACDALRLPADAGQQQKLLRYIQMQRW
NRTYNTATRDQGLVQHLFDSLSVSLVPLERGLPAAGSGARVLFVGGGGLPGVY
LAIMRAHWDVTCVDAVEKKTAFVRQAGALGLEPQLQAHTRIEQLPEAQCDVVISRA
ASLQDFAKLAGRHVRREGGTLVAMKGVDPDEIQALQQHGHVTEPLVVPALDAQR
CLIMWERSQNI"
1917..2135
/gene="gidB"
/locus_tag="BPP0002"
/notes="PS00430 TonB-dependent receptor proteins signature
1."
2004..2579
/gene="gidB"
```

```

/misc_feature
/locus_tag="BPP00002"
/notes="HMPfam hit to PF02527, Glucose inhibited division
protein, score 1.7e-50"
2606. 3403
/gene="para"
/locus_tag="BPP00003"
/notes="synonym: soj"
2606. 3403
/gene="para"
/locus_tag="BPP00003"
/notes="ortholog of Bordetella pertussis (BX470248) BP0003
Similar to Caulobacter crescentus chromosome partitioning
protein ParA or Cc3753 SWALL:PARA CAUCR (SWALL:O05189)
(267 aa) fasta scores: E(): 9.6e-46, 49.41% id in 255 aa,
and to Bacillus subtilis sporulation initiation inhibitor
protein Soj SWALL:SOJ RACSJ (SWALL:P37522) (253 aa) fasta
scores: E(): 1.3e-43, 48.98% id in 247 aa"
/codon_start=1
/transl_table=11
/product="Para family protein"
/protein_id="CAE39744.1"
/db_xref="GI:33564802"
/db_xref="UniProt/TREMBL:Q7W2H9"
/translation="MKNIIPPSARVFCIANQKGVGKTTTAINLAAGLATHKQRVLL
VDLPQGNATMGSGIDKSTLESNLYQLVGEAGIEQTVRSSEGGYDVLPAFRELSGA
EIDLVDQDERERQLKAAIDKAGEYDFVILDCPTLSLLTLNGLAAHGVIIIPMQCEY
FALGSLDLVNTIKRVHRNINNELRVIGLLRVMPDPMTLQQQVSAQLESFHDGKVFT
TVVPRNVRLAEASVGMPCGVVYDRASRGAQAYIAFGAEMIERVKELD"
2879. 3214
/gene="para"
/locus_tag="BPP00003"
/notes="HMPfam hit to PF00991, Para family ATPase, score
1.5e-33"
3410. 4216
/locus_tag="BPP00004"
/locus_tag="BPP00004"
/locus_tag="BPP00004"
/notes="ortholog of Bordetella pertussis (BX470248) BP0004
Similar to Caulobacter crescentus acetyltransferase, GnaT
family Cc2209 SWALL:Q9A685 (EMBL:AE005892) (207 aa) fasta
scores: E(): 1.5e-11, 34.63% id in 205 aa, and to
Rhizobium meliloti putative acetyltransferase protein
R02624 or Smc02449 SWALL:Q92M32 (EMBL:AL591791) (308 aa)
fasta scores: E(): 5.5e-10, 31.42% id in 245 aa"
/codon_start=1
/transl_table=11
/product="Putative acetyltransferase, GnaT family"
/protein_id="CAE39745.1"
/db_xref="GI:33564803"
/db_xref="UniProt/TREMBL:Q7W2H8"
/translation="MPAVLAVLAHPAPAAGGELARLEEAALNATTVREOMLYEGWLVL
RWAPSRKARCVNVLGLSTPLDERLAYCTRWYAGHVLIVRLTSGIDPFLSDAQL
HERGFVAFRTCMAPVPVPTAQASGLVFERTTGPRAQENGLRGHDDATAEHA
ARLQAAPLENPLDVLADSGECIAAGLAADFVGLGFDIVTAPQRRRCGGAVLKY
LUEQASHAGARQAYLQVEFQNTAARALYARFGEVDRTYTYRMSSEATPV"
3929. 4165
/locus_tag="BPP00004"
/notes="HMPfam hit to PF00583, Acetyltransferase (GNAT)
family, score 2.5e-14"
4260. 5177
/gene="parB"
/locus_tag="BPP00005"
4260. 5177
/gene="parB"
/locus_tag="BPP00005"
/notes="ortholog of Bordetella pertussis (BX470248) BP0005
Similar to Pseudomonas putida probable chromosome
partitioning protein ParB SWALL:PARB PSEPU (SWALL:P31857)
(290 aa) fasta scores: E(): 1.3e-40, 48.81% id in 295 aa"
/codon_start=1
/transl_table=11
/product="probable chromosome partitioning protein ParB"
/protein_id="CAE39746.1"

```

```

/misc_feature
/db_xref="GI:33564804"
/translation="MATKKPKGLRGDLADGADVPADINIGKAPAAPEGPATLPV
SKWRAGKYQPTFMDGALNEIAESIRTOIMQPIVLRALGADAPQGVETIAGERRFR
AAQAGLKVFPVLVREVADENAAVVALIENIQREDLNPLBEAHGVRLDDEGLTHEQ
AAQATGRSRSATNLLRLNLNLAAPVOTMILLAGDVMGHARALLVDAATOIQIANQVI
AKRLSVREAKLVARAQKVDVDAAPRKKGNGASRDVTRLEALSDDLGLTRVALKVARE
KGQIVIDFHGWBHLNLSLLRQGLSGVVDA"
4377. 4658
/gene="parB"
/locus_tag="BPP00005"
/notes="HMPfam hit to PF02195, ParB-like nuclease domain,
score 5.9e-34"
5170. 6171
/gene="ansB"
/locus_tag="BPP00006"
5170. 6171
/gene="ansB"
/locus_tag="BPP00006"
/EC_number="3.5.1.1"
Query Match 5.3%; Score 63; DB 1; Length 348251;
Best Local Similarity 48.5%; Pred. No. 8e-08;
Matches 174; Conservative 0; Mismatches 185; Indels 0; Gaps 0;
QY 276 CTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGGCAGTGGC 335
DB 228866 CTGGCGCGCGCGCCCATCGCGCGCAACGCGGATCACCAGCAGCGCTCCAGCGCC 228807
QY 336 CGGTTCATGACAGACATCCCGGAGGTAAAGCGTCCCTGTGTCACTTCTTCTCAATG 395
DB 228806 TGCGGCCAGCGCGCGGTTCGGACGCGGCACAGCAAGCGGTTCGCGCATGCGCACG 228747
QY 396 ACATCAGGGATCCCGCGCTCTCCTCTGCGGATAGCGGCACGCGGAGACTGACGCTTCA 455
DB 228746 ACTTCGGCAGCGCGCGCCACGTTGCGCGGATCACCAGCAGCGCGCGCGGCGCTCG 228687
QY 456 GCCAGTACCATCAACAAACGCTTTCATTTCCGAGGATGACCAACCACTGGCAATCCCG 515
DB 228686 ACCATGGCCCAACCCAGGCGCTCGTTCGGGTTCGCGCGGCGAACAACATCGAAAGCGCC 228627
QY 516 TAGACGGTAACTGGGAAAAGGCACTGCCATTAACACATCTCGGTCAATCCAGG 575
DB 228626 AGCACATTTGGGATGTCGTCGGATGCCCCGTGAGTGACGCGCGCGCCAGCCCGCAGG 228567
QY 576 TGTTCTGTCTGTCGACGACGCTGCTTCTGTTATTTTCACGCGCGCGCCACACAG 634
DB 228566 CGTTCGCGCGCTTCGCGCAGCAGGCGCGCGGATCGCCCGCGCTCGCCGACGACAG 228508
RESULT 24
BX640437/c
LOCUS
DEFINITION
BX640437
BX640437 BX470250
VERSION
BX640437.1 GI:33575039
KEYWORDS
complete genome.
SOURCE
Bordetella bronchiseptica RB50
ORGANISM
Bordetella bronchiseptica RB50
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Bordetella.
REFERENCE
1
AUTHORS
Parkhill,J., Sebaihia,M., Preston,A., Murphy,L.D., Thomson,N.,
Harris,D.E., Holden,M.T.G., Churcher,C.R., Bentley,S.D., Harris,B.,
Mungall,K.U., Cerdeno-Tarraga,A.M., Temple,L., James,K., Harris,B.,
Quail,M.A., Achtman,M., Atkin,R., Baker,S., Basham,D., Bason,N.,
Cherevach,I., Chillingworth,T., Collins,M., Cronin,A., Davis,P.,
Doggett,J., Feltwell,T., Goble,A., Hamlin,N., Hauser,H.,
Holroyd,S., Jellings,K., Leather,S., Moule,S., Norberczak,H.,
O'Neil,S., Ormond,D., Price,C., Rabinowitsch,E., Rutter,S.,
Sanders,M., Saunders,D., Seeger,K., Sharp,S., Simmonds,M.,
Skelton,J., Squares,K., Stevens,K., Unwin,L.,

```



```
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="DNA transport competence protein"
/protein_id="AAO29400.1"
/db_xref="GI:28057552"
/transl_table="MKSFMRLKSLLLFLVLAFTVHAANKVDINTASAEWDKVLVNI
GPSKASAIKYTREENGPPKSVBEIALVKGIMKTVERNRDLIEIGTRMAPAKHAKGTE
LKSVMGR"
2603..3625
/locus_tag="PD1559"
2603..3625
/locus_tag="PD1559"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO29401.1"
/db_xref="GI:28057553"
/transl_table="MTSSDFSDTTRATQRLQWIRTTLADSTATLQRASSDAGFRSYWR
STNGLSHLMADPHELENPHQWIRIHALITSGGVRVPHILAQDLAEGFLLEDLGIPT
TLAQRDLADNADTLPDAALQIALQICIVPNDLPRENTALLERDASLFDWFLYRHL
NLARTDLALQVQQQMDNALTOPRLVHRDFMERNLMLTDDGVYLDFOQCTVG
PVAYDPSLFRDYSWPLARVDRKWLTHYHARANAAKIPVQTLPHFLRDADMGVQRH
LKNIGIFARLHYDGRKSYLENIPRFISYIEEILPRHPHTLAPLAEILIEHRIKPAALR
MITEST"
3622..4344
/locus_tag="PD1560"
3622..4344
/locus_tag="PD1560"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="nucleotidyl transferase"
/protein_id="AAO29402.1"
/db_xref="GI:28057554"
/transl_table="MKALIFAAGIGQRMRLTYNTPKPLCAGGEPLIVNLRKLAAL
GISVNTATWLSQFPEILGDSQRFGLRFLYSNEGSLPLETGGMLHALPLGNAPF
LAINGDIWTDLTRLPEVGDHMLVNNPEVHPQGDVFLQADSVLDRTHGPIPL
TFAGLGYRSQLADWENIIGDTPDTHAQPRPKLPLLRAMRSRGIHGTTHRGQWT
DVGTQRLHALDWTLSPEARF"
complement(4527..5228)
/locus_tag="PD1561"
complement(4527..5228)
/locus_tag="PD1561"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="replication related protein"
/protein_id="AAO29403.1"
/db_xref="GI:28057555"
/transl_table="MSVSQPLALRYSSDQRFETYLHAPSGLIAQLRALVDQCSREWI
YLVGASGKTHALFALCAAEGQSVAYLPQVAVGRRLDALEALEGCVLYVLGL
QAIYGMREDIALPFDNRAAGITLLYTAIRAIIDGLTLPDLRSRVAQCIRIAPL
TLDDVGRASVLRERARQRRGLMLDEAIEWLLSRDRLDGLVMLLERDLRESLATRR
VTIPFLRVLAEQGG"
complement(5225..6382)
/locus_tag="PD1562"
complement(5225..6382)
/locus_tag="PD1562"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="permease"
/protein_id="AAO29404.1"
/db_xref="GI:28057556"
```

```
/translation="MMFSPVEIARFLRLKRWYLVGGVLMMAWLLTPVLIPFVLSVP
LGMADPLRVKVEWGVSRMAWCVFFVLLVLLVLMIFVFLIERQVVLINTFPQ
MHDVVNTVIPWEQTSQRLTDEQIMQWHSNWEQAGGVARGFGVSRSGFV
MVTWNLALLPILSFYFLDWDKWVERVAAAI PRAYIGVSRLATANEVIAFGFV
QFLVNLALGAIYAGLSVGLRLGLIAGLISFIPYLGAITGVVLAIATVQVH
GLDQLLVGVVGVGQQLLESVLTPTIVGDKIGLHPVAVIFSVMAAGGLFGPVGML
LALPAAVNVNLLRYAHERVIESEFYKQYLAVLALHECKSDENCDCGKKDAV"
complement(6379..7551)
/locus_tag="PD1563"
complement(6379..7551)
/locus_tag="PD1563"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAO29405.1"
/db_xref="GI:28057557"
/transl_table="MAAMPYFGQSLFCISHAMELSMRVFFLPFIALCAVLPPVPAVA
ORAPHIQSDVAKESQSIYEAEVPVNSQSDSRKAGIARALSIVLGKVTGDRSRNAKLRP
EVSQALDAAAFVDSYDROQSSPGGAPTSTMLTHTVHFRPDEVNALGALALPLMP
QRPKPKVMLAIDDGSPRLVGVQQRVARSVLDRAITERGRLGPRBEEQALVTA
IMRSDVASVLSASTRYTSSMLVGLKYSNGNWAADI FVDGGEQVASWSSNADALR
VMADGADGAADLVKRYAKVPLTGTGTVYGVINGIRSDADYLRVSAALQRVPVVRSM
IPKASAERLVYLDLMTGITGLNRLMGDSALQPVAATLESQSGELVEYVLK"
7602..8657
/locus_tag="PD1564"
complement(8657..9811)
/locus_tag="PD1564"
/note="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="5'-phosphoribosyl-5-aminimidazole synthetase"
/protein_id="AAO29406.1"
/db_xref="GI:28057558"
/transl_table="MTTPPPSTPLTYRDAGVDIDAGNALVERIKPLVTRTFPEVM
SGLGGFALFNLAGTKYKPEVLVSGTQVGTGKLAQOLNRHTTIGDLVAMCVNDLV
QGAEPFLFDYFATGKLDITATAVISGIALGCEQSCALIGGETABMDPMYPPGEYD
LAGFCVAAVEKSQLDSQVREDDVLIGIASGPHSNGYSLIRIYERAGSPALDIDH
GTRLIDTLNAPTALYKPKILKLLHSDAIHANAHITGGGILTEIIIRVIPNLGLRID
ANAWTPPVFWLQREGALADTEMWRTFNCGIGFVLVATPNQVAPLQALDQGLAHW
QIGRVVTPVDNERVHIG"
Query Match 5.1% Score 60.2; DB 1; Length 300885;
Best Local Similarity 51.7%; Pred. No. 6.4e-07;
Matches 137; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
Qy 285 GCTGCCATCATCCGGAAGCGTTCGGTCCGGATAAAAAATCGGCAGTCGCCCGTCCAT 344
Db 298291 GCTGCCCATCATCGGCCCGGAGTTCGGATCACAAATCGCAATCCGCCAC 298350
Qy 345 GCAGACACATCCCCACCGGTAAACAGGTCCTCTGTACATCTTCTGAATGACATCAGG 404
Db 298351 GCAGGACAGTCCCGCGGCGGAGCAGCATCCCGTCTGTACATCTCTCCGCA 298410
Qy 405 ATCCGCGCGTCTCTACCTGGCGATAACGGGCACGCCGAGACTGACGCTTCAGCCAGTACC 464
Db 298411 GCACACCCACATCATCTCAACGACAGGCACACCGGATGCTTCGGGCTCTCGGAGACA 298470
Qy 465 ATACCAAAACGCTTTCATTTCCGAAGGCATGACCACCACTGGCAATCCGCTAGACCGGT 524
Db 298471 CGTCCAAACGCTCTCCGCGACAAAAGAGGCATCGCCAGCATCGCCACGCTCGAAATAA 298530
Qy 525 AACGCTGGGAAAAGGCGACCTGCCA 549
Db 298531 CGCGCGCGGTCAGTCAACCCCAACCCA 298555
RESULT 26
```


/product="hypothetical protein"
/protein_id="AAF83686.1"
/db_xref="GI:9105788"
/translation="MELTAIFPQSNHDVYAVLAHCSAYRARCIAHFTGSGVSGPELHLE
LVLLMAKQRAQHNAVCRLPKIEHTDQREALAVQGSVKVRRESERKNQRLCLIDY
"

gene

4321..4635

/locus_tag="XF0877"

4321..4635

/locus_tag="XF0877"

/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1

/transl_table=11

/product="hypothetical protein"

/protein_id="AAF83687.1"

/db_xref="GI:9105789"

/translation="MLGVLCRRRRRFTCGHOTFGQFVQRCGLLFQMGEESSRWQVWHIM
SPNLISATSKFQANLFGPGYQLVFECVLPQWQRRARIHLLVFNHGGEGRRPKN
IG"

gene

4628..5401

/locus_tag="XF0878"

4628..5401

/locus_tag="XF0878"

/note="similar to GI|1783247 (percent identity: 29 %/query
alignment coverage: 84.4 %/subject alignment coverage:
77.8 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"

/codon_start=1

/transl_table=11

/product="conserved hypothetical protein"

/protein_id="AAF83688.1"

/db_xref="GI:9105790"

/translation="MGSMHFGQAAEYGRGLFMATPILMYHIAKVPKQVRLRLGLV
TPAPARQMLLRLGLSCLSMAAMPYLRGSGKVMVVTLDGGLDNLQALPLVQ
AHGSACTVYVSSGLARENTDAERLNVCKPLMSPAOVROWHDAGMEVGAHTRSHPL
SGTAAQHEIACGRDLEQCIGAPVQCYPYGVDTTPVIDAVCDAGYAAATTRR
GRVFPQGLWTLPRVPSYRHLFPQFALRTLTGYEDRI"

gene

5398..6510

/locus_tag="XF0879"

5398..6510

/locus_tag="XF0879"

/note="similar to GI|3337079 (percent identity: 26 %/query
alignment coverage: 94.9 %/subject alignment coverage:
95.4 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"

/codon_start=1

/transl_table=11

/product="lipopolysaccharide biosynthesis protein"

/protein_id="AAF83689.1"

/db_xref="GI:9105791"

/translation="MKVLFVGTSGGGGASHFVGLVRAMAETNHQSTVLVHPDGLIA
RLOQOTAIRLFTAFPRVMDLRLGLYALRALRLVHPDVLGDFGKEYWPLLMGRLYR
LPVLPRHLRPLPMNCFSTYVWPLADFEAVSAYARHLYLAEGMPPERVOLVNPVDT
DALRPDRVRAMLHELGNDEDVLVCGFRHKGKGVFVLAEMEQAMQEBRLCCL

Query Match

Best Local Similarity 5.0%; Score 59.2; DB 1; Length 13334;

Matches 130; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

Qy 285 GCTGCCATATCCGGAAGCTTCGGTCCGGATAAAAATCGGCAGTGGCCGGTCCAT 344

Db 6395 GCTCTGCCATCCGGCCCGGGGTTGCGGATACAAAAATCGCAATCCGATCCCGCCAC 6336

Qy 345 GCACACATCCCCACGGGTAAACAGCGTCCCTGTCACTTTCTGAATGACATCAGGG 404

Db 6335 GCAGGACGTTCCCGGGGCGGACGAGATCCCGGTGTACCATCTTGCATGTCTCCGCA 6276

Qy 405 ATCCCGCCCGTCTCACTGGGATAAACGGGACCGCGGAGACTGACGTTTCAGCCAGTACC 464

Db 6275 GCCCCACCCACATCACTCACAAGCACAGGCACACCGGATGTTGCGCTCTTCGGGAGACA 6216

Qy 465 ATACCAACGGTTCATTTTCGAAGGATGACACACACATGGCAATCCGGTAGACCGGT 524

Db 6215 CGTCCAAACGCTCTCCGGCAACAAAGAGGATCGCCAGATCGCCAAACGCTGGAATAA 6156

Qy 525 AACGCTGG 532

Db 6155 CGGCGCGG 6148

RESULT 27

AE000516 24

WPCOMMENT

Sequence split into 44 fragments LOCUS AE000516 Accession AE000516

Fragment Name	Begin	End
AE000516_00	1	110000
AE000516_01	100001	210000
AE000516_02	200001	310000
AE000516_03	300001	410000
AE000516_04	400001	510000
AE000516_05	500001	610000
AE000516_06	600001	710000
AE000516_07	700001	810000
AE000516_08	800001	910000
AE000516_09	900001	1010000
AE000516_10	1000001	1110000
AE000516_11	1100001	1210000
AE000516_12	1200001	1310000
AE000516_13	1300001	1410000
AE000516_14	1400001	1510000
AE000516_15	1500001	1610000
AE000516_16	1600001	1710000
AE000516_17	1700001	1810000
AE000516_18	1800001	1910000
AE000516_19	1900001	2010000
AE000516_20	2000001	2110000
AE000516_21	2100001	2210000
AE000516_22	2200001	2310000
AE000516_23	2300001	2410000
AE000516_24	2400001	2510000
AE000516_25	2500001	2610000
AE000516_26	2600001	2710000
AE000516_27	2700001	2810000
AE000516_28	2800001	2910000
AE000516_29	2900001	3010000
AE000516_30	3000001	3110000
AE000516_31	3100001	3210000
AE000516_32	3200001	3310000
AE000516_33	3300001	3410000
AE000516_34	3400001	3510000
AE000516_35	3500001	3610000
AE000516_36	3600001	3710000
AE000516_37	3700001	3810000
AE000516_38	3800001	3910000
AE000516_39	3900001	4010000
AE000516_40	4000001	4110000
AE000516_41	4100001	4210000
AE000516_42	4200001	4310000
AE000516_43	4300001	4403837

Continuation (25 of 44) of AE000516 from base 2400001 (AE000516 Mycobacterium tuberculosis

Query Match 4.8%; Score 57.2; DB 1; Length 110000;

Best Local Similarity 54.2%; Pred. No. 4.9e-05;

Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 272 TTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGGTCCGGATAAAAAATCGCGCAG 331

Db 47408 TTCTCGCGCGCGCGCCCATCGCAGCGGCCGATCCCGATCGATCAGCAACTCGGCCAC 47467

Qy 332 TGGCGCGGTCCATGCGAGACATCCGCCACGGGTAAACAGCGTCCCTGTCACTTCTCTG 391

Db 47468 GGGCTCGGCCACCGCGTCCACCGACCTACCGTACCGTCTTGTGTGTG 47527

Qy 392 AATGACATCAGGATCCCGCCCTCTCACTGGCGATACCGGCACCGGAGACTGAGGC 451

hypothetical 23.2 kd protein in sbm-fba intergenic region
(211 aa), FASTA scores: opt: 379, E(): 1.5e-19, (37.3%
identity in 185 aa overlap); and Q11154|rv0488
HYPOTHETICAL 20.9 KD PROTEIN from Mycobacterium
tuberculosis (201 aa), FASTA scores: opt: 784, E(): 0,
(63.4% identity in 186 aa overlap). BELONGS TO THE
LYSE/YGGA FAMILY."
/codon_start=1
/transl_table=11
/product="PROBABLE CONSERVED INTEGRAL MEMBRANE PROTEIN"
/protein_id="CAD96861.1"
/db_xref="GI:31618757"
/db_xref="GOA:Q10871"
/db_xref="UniProt/Swiss-ProT:Q10871"
/translation="MNSPLVVGFLACFTLIAIGAQNAFVLQRQHREHLVPVALCT
VSDTLTAAGTAGCAATGAPRALNVVKFGAAELIGYGLIAARRARPVALIPSGA
TPVRLAEVLNVCFAAPTFLPNHYLDTVVLLGALANEHSQDWLFGLGAVTASAVWFAT
LGFGAGRLGLTFPWSRWIIDLGLIAMVMVALGISLTVT"
4942..5370
/locus_tag="Mb2009"
/locus_number="3.2.1.14"
/ec_number="3.2.1.14"
/note="Mb2009, -, len: 142 aa. Equivalent to Rv1987, len:
142 aa, from Mycobacterium tuberculosis strain H37Rv,
(100.0% identity in 142 aa overlap). Possible chitinase
(EC 3.2.1.14), similar to several e.g. P36909[CHIT STRLI
chitinase c precursor (619 aa) FASTA scores, opt: 324,
E(): 1.2e-14, (39.5% identity in 129 aa overlap)."
/codon_start=1
/transl_table=11
/product="POSSIBLE CHITTINASE"
/protein_id="CAD96862.1"
/db_xref="GI:31618758"
/db_xref="GOA:Q10870"
/db_xref="UniProt/Swiss-ProT:Q10870"
/translation="MAGLNIVRRWRTAHATVSALIVAILGLAITPVASAATARATL
SVTSWTWGTFARTISNGSTAPLTDWKLEDFLPAGESVLHTMNSTARSGETHYVLSP
ANNNRIIAPGSGATGRLRGSLTGTSYPPSSCLLNGQPCT"
5596..6135
/locus_tag="Mb2010"
/locus_number="3.2.1.14"
/ec_number="3.2.1.14"
/note="Mb2010, -, len: 179 aa. Equivalent to Rv1988, len:
179 aa, from Mycobacterium tuberculosis strain H37Rv,
(100.0% identity in 179 aa overlap). Probable
methyltransferase (EC 2.1.1.-), similar to
ERME SACER[P07287 rna adenine n-6-methyltransferase (370
aa), FASTA scores: opt: 259, E(): 2e-11, (35.1% identity
in 171 aa overlap); contains PS00092 N-6 Adenine-specific
DNA methylases signature. Also similar to Mycobacterium
tuberculosis Rv1010 ksgA 16S rRNA dimethyltransferase."
/codon_start=1
/transl_table=11
/product="PROBABLE METHYLTRANSFERASE"
/protein_id="CAD96863.1"
/db_xref="GI:31618759"
/db_xref="UniProt/TREMBL:Q7TZ64"
/translation="MSALGRSRRAWHRLHDEWAARVVSAAVRPBELVFDIGAGEG
ALT AHLVRAGARVAVLHPRVGVLRFRFPGITVVHADAAASIRLPGPRFVYANPPY
GISRSRLRTLAPNSGLVAADLVQLRALVCKFASRNARRFTLTVLGLMLPRRAFLLPPP
VDSAVLVRRKCGDQGR"

	Query Match	Best Local Similarity	4.8%; Score 57.2; DB 1; Length 306050;
	Matches 116; Conservative	0; Mismatches 98; Indels 0; Gaps 0;	
Qy	272	TTCCTCGCGTGGCTGCCCATCATCCGAGCGTTCCGTCGGGTATAAAAATCGGCAG	331
Db	224107	TTCCTCGCGCGCGCGCCATTCGCGAGCGCCCGATCCCGATCAGCAACTCGGCAC	224166
Qy	332	TGCGCCCGTCCATGCAGACACATCCCCACCGGTAACAGCGTCCCTGTGCATATTCTTCG	391

misc_feature	<pre>/gene="pbpB" /locus_tag="Rv2163c" /function="Involved in peptidoglycan biosynthesis." /note="Rv2163c, (MTCY270.05), len: 679 aa. Probable pbpB, penicillin-binding membrane protein, similar to many bacterial PBP2 proteins e.g. P11882 PBP2_NEIME PENA NMA2072 NMB0413 penicillin-binding protein 2 (pbp-2) from Neisseria meningitidis (serogroups A and B) (581 aa), FASTA scores: opt: 665, E(): 1.6e-31, (33.2% identity in 591 aa overlap); etc. Also similar to Rv0016c and Rv2864c from Mycobacterium tuberculosis (2.8e-10). Contains P500017 possible ATP/GTP-binding site motif A (P-loop) near C-terminus. FASTA best: PBP2_NEIME P11882 penicillin-binding protein 2 (pbp-2). (581 aa) opt: 665, E(): 1.6e-31; (33.2% identity in 591 aa overlap)" /codon_start=1 /transl_table=11 /product="Probable penicillin-binding membrane protein pbpB" /protein_id="CAB08664.1" /db_xref="GI:2104315" /translations="MSRAAPRRASQSQSTRPARGLRPPGCAQEVGQRRPKQTKARQ AQATKSRPATRSDVAPAGSTRARRTRQVVDVCTRGASFVFRHRTGNVILVLMVA ATQLFLQVSHAARQAAGQLKVDVQPAAGSIKVDNRNDRLAFTIEARALTPOPK RIRQLEARKKTSAAAPDPQQLRDIAQEVAGKLNKPDAAAVKKLQSDETFFYLAR AVDPVASAI CAKYPEVGAERDLQVPGGSLAANVGGIDWDHGLGLEDLSLDAVL AGTDSVTVDRGSDGVVIGSYRNKHAVHGSTVLTLDNDIQFYQQQVQQAKNLSG AHNVSAVLDAKTEVLAMANDTDFDPSQDIGRQDKLGNPAVSPFSPGSGVNVKIVA ASAVIHGLSSPDEVILQVPGSIQMGVTVHDHWEHGVPTTGTGFGKSNVGTMLLS QRVGPRYDMLRFGQRTQVGLPGESAGLVFPIDQWSGTFANLP1QGGLSMTLL QMTGMQAANDGVVPPRIIKATVAPDGSRTPEPRDDIRVVSAQTAVQKMLRAV QVDRPMGYQGTGTGVPGYQMGKTGTAAQINPGCGVFDVYVITFAGIATADNP RVVIGTMDLNPARNSDGAGHSAAPLFHNIAGLMQRENVPISDPGPPVLQAT" /complement(470. .493) /gene="pbpB" /locus_tag="Rv2163c" /note="P500017 ATP/GTP-binding site motif A (P-loop)" /complement(2245. .3399) /locus_tag="Rv2164c" /complement(2245. .3399) /locus_tag="Rv2164c" /function="UNKNOWN" /note="Rv2164c, (MTCY270.04), len: 384 aa. Probable pro-rich conserved membrane protein, equivalent to ML0907 AL022602 putative conserved membrane protein from Mycobacterium leprae (377 aa) (AL022602). FASTA scores: opt: 1495, E(): 1.7e-56, (62.217% identity in 397 aa overlap)." /codon_start=1 /transl_table=11 /product="PROBABLE CONSERVED PROLINE RICH MEMBRANE PROTEIN" /protein_id="CAB08663.1" /db_xref="GI:2104314" /translations="MRAKREAPKSRSDRRRRADSPAAATRTTNSAPSRIRSRAG KTSAPGRVSRPGQTSFMLSFPDPAPAKNTSQAKAKAKAKAPKLVRPTPME RLAAKTSIDLRLTANKVPFVVLVIGSLVGLGLTLWSTDAAERSYQLSNARENT RMLQKHLEARDREAAAPALAAEARQGMIPTRDTAHLVQDPDGNMWWVGTTPKA DGVPPEPLTKLEDPPPPKPAAPLEVPRVPTPGDDPAPPARSGPEVLRTPDGT ALTGCAHLPTQAGLPQGPVPIPGAPGMPAPPLGAVPSPAPENPVPVLQGAAPPA GLGPAVPAATPGLSGSQPMVAPPAPVPANGEQFGVTAFTAPGAPR" /complement(3396. .4586) /locus_tag="Rv2165c" /complement(3396. .4586) /locus_tag="Rv2165c" /function="UNKNOWN" /note="Rv2165c, (MTCY270.03), len: 396 aa. Conserved hypothetical protein; shows strong similarity to several hypothetical bacterial proteins but has extra 80 aa residues at N-terminus FASTA best: YLXA_BACSU Q07876 hypothetical 35.3 kDa protein in ftsL (311 aa) opt: 781, E(): 0; (45.6% identity in 296 aa overlap), BELONGS TO THE YABC (E.COLI), YLXA (B.SUBTILIS) FAMILY"</pre>		gene	<pre>Query Match 4.8%; Score 57.2; DB 1; Length 348247; Best Local Similarity 54.2%; Pred. No. 6.3e-06; Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;</pre>	
	<pre>/codon_start=1 /transl_table=11 /product="CONSERVED HYPOTHETICAL PROTEIN" /protein_id="CAB08662.1" /db_xref="GI:2104313" /translations="MQTRAPVSLPEATLAFNPNRVSSDRDLGAGAFCIAASRSTA CWTGGITVADPGSGPTGFGHVPVLAQRCFELLTPALTTRYPDGSOAVLLDATIGAGG HAERFLGELGLRLIGLDPTALDVARSRLVRFADRLTLVHTRYDCLGAALAESGYA AVGVDDGLFDIGVSSMQLDRAERGFAYATDAPLDMRMDPTTLPATAADVNTVDEAAL ADILRRYGERFARRIAGIVRRRAKTPPTTAELVALLYQALPAARVGGHPAKRT FQALRIANDELESURTAVPAALDAIGAIGRIANLAYQSLERIVKRVFAEAVASATP AGLPFLPGHPRFSLTHGAERASVAEIERNPRSTFVRLRLQRVHEHQSOQWATE KGS"</pre>			<pre>complement(4588. .5019) /locus_tag="Rv2166c" /complement(4588. .5019) /locus_tag="Rv2166c" /function="UNKNOWN" /notes="Rv2166c, (MTCY270.02), len: 143 aa. Conserved hypothetical protein; shows strong similarity to several hypothetical bacterial proteins such as YLLB_BACSU P55343. Is equivalent to Mycobacterium leprae hypothetical protein ML0905 (143 aa, 92% identity) MLCB268.11c >ep 069561 YL66 MYCLE HYPOTHETICAL 16.1 KDA PROTEIN ML0905 >gi 3080482 emb CAA18677.1 (AL022602) >gi 13092975 emb CAC31286.1 (AL583920). FASTA scores: ML0905 ML0905 conserved hypothetical protein (143 aa) opt: 873, E(): 3.1e-52; 92.254% identity in 142 aa overlap; YLLB_BACSU P55343 hypothetical 16.6 kDa protein (143 aa) opt: 340, E(): 3.6e-17; (35.0% identity in 143 aa overlap) . BELONGS TO THE YABB (E.COLI), YLLB (B.SUBTILIS), MG321 (M.GENITALIUM) FAMILY" /codon_start=1 /transl_table=11 /product="CONSERVED HYPOTHETICAL PROTEIN" /protein_id="CAB08661.1" /db_xref="GI:2104312" /translations="MPLGTYTPKLDDKGRLLTLPKPRDALAGLMVTKSQDHSLAYYP RAAFEQLARRASKAPRNPPEARAFNLNLAAGTDEQHPDSQGRITLSADHRRVYASLKD CVVIGADVLEIWDQAQWQYQIHEENFSAASDEALGDF" /complement(5278. .6632) /insertion_seq="IS6110-6, len: 1355 bp. Insertion sequence IS6110." /notes="IS6110-6, len: 1355 bp. Insertion sequence IS6110." /complement(5278. .5305) /notes="28 bp Inverted repeat at the left end of IS6110; GAGTCCCGGACTCACCGGCGGTCA" /complement(5320. .6360) /locus_tag="Rv2167c" /complement(5320. .6360) /locus_tag="Rv2167c" /function="REQUIRED FOR THE TRANSPOSITION OF THE INSERTION ELEMENT IS6110." /notes="Rv2167c, (MTCY270.01), len: 346 aa. Probable IS6110 transposase. FASTA best: TRAY MYCTU P19774 putative transposase for insertion sequence (identical)" /codon_start=1 /transl_table=11 /product="PROBABLE TRANSPOSASE" /protein_id="CAA17494.1" /db_xref="GI:3242294" /translations="AEALAAQORRIAKGRDPKRVFLGRARPASTLITRFIADHQ GHREPGDLRWGVESICTQLTELGVPIAPSTYYDINREPSRELDRGELKHSRVH AANYGVYGKAVLTLNREGIIVARCTVERLMTKLGLSGTTRGKARTTIADPTARP ADLVORRFGPPAPNRLWADLVTVSTWAGFAVAVFTDAYARRILGWRVASTWTGSMV LDALISQAIWTRQEGVLDLKDVIHHTDRGSQVTSIRFSELAENGQPSVAGVGSYD NALAEITNGLYTKELIKPGKPNRSIEDVELATRWDFNHRRLYQYCGDVPVPELEA AYAQRQRPAAG" /complement(6255. .6581)</pre>	


```

Qy 272 TTCTCTGGCGTGGTCCATCATCCGAGCGTTCCGGTCGGGATAAATAATCGCGAG 331
Db 25235 TTCTCGCGCGCGCGCCATCGGACGCGCCCGATCCGATCGATCGAATCGGCAC 25294
Qy 332 TGCGCCGCTCCATGACACATACCCACCGGTAAACAGCGTCCCTGTGACATTTCTCTG 391
Db 25295 GGCCTGCGCCACCGCGTCCACCGACCTACCGTCGACCACTAGCCAGTCTTGTGTCTG 25354
Qy 392 AATGACATCAGGATCCGCGCGCTCTCATCTGGCGATACCGGACGCGGAGACTGACGC 451
Db 25355 CACCGTTTCGCGCGCTCCGCGAGAAATTCGCGCGATTAACCGGACGCGCGCGGAGGC 25414
Qy 452 TTCAGCAGTAGTACCATACCAACGCTTCATTTTCC 485
Db 25415 TTCAGGACACGATGATGCCAAGCCTCGACGTCC 25448

RESULT 30
AF498417/c
LOCUS
DEFINITION Pseudomonas aeruginosa serotype 06 putative O-antigen biosynthesis
AF498417 AC104736
VERSION
KEYWORDS
SOURCE
ORGANISM
Pseudomonas aeruginosa
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 14734)
Raymond, C.K., Sims, E.H., Kas, A., Spencer, D.H., Kutayavin, T.V.,
Ivey, R.G., Zhou, Y., Kaul, R., Clendinning, J.B. and Olson, M.V.
Genetic variation at the O-antigen biosynthetic locus in
Pseudomonas aeruginosa
J. Bacteriol. 184 (13), 3614-3622 (2002)
JOURNAL
MEDLINE
22053227
PUBMED
12057956
REFERENCE
2 (bases 1 to 14734)
Raymond, C.K., Sims, E.H., Kas, A., Spencer, D.H., Kutayavin, T.V.,
Ivey, R.G., Zhou, Y., Kaul, R., Clendinning, J.B. and Olson, M.V.
Direct Submission
TITLE
Submitted (04-APR-2002) Genome Center, University of Washington,
Box 352145, Seattle, WA 98105-2145, USA
JOURNAL
On May 13, 2002 this sequence version replaced gi:17975265.
COMMENT
FEATURES
source
1. 14734
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/serotype="06"
/db_xref="taxon:287"
<1..104
/feature="ORF 1; C-terminal coding region of rpsA"
/codon_start=3
/transl_table=11
/product="RpsA"
/protein_id="AA027812.1"
/db_xref="GI:20560067"
/translation="EKDAMKELRKQEVESAGPTTIGDLIRAQMNQGS"
241..525
/feature="ORF 2; hmd/ihfB"
/codon_start=1
/transl_table=11
/product="Hmd"
/protein_id="AA027813.1"
/db_xref="GI:20560068"
/translation="MTKSELIERIVTHQGLSADVELAKTKMLEQMSQALATGDRIE
551..805
/feature="ORF 3; hypothetical orfA"
/codon_start=1
/transl_table=11
/protein_id="AA027814.1"
/db_xref="GI:20560069"

```

```

/translation="MOMLGRILTAVFVLVVALVVLVLENQNSVLSLAFGLMHTPDGP
LSLPLVSLATILGSLGSLGALLRLRGRROKSLRSSTSL"
1190..2254
/feature="ORF 4; wzz; similar to chain length determinant
protein"
/codon_start=1
/transl_table=11
/product="Wzz"
/protein_id="AA027815.1"
/db_xref="GI:20560070"
/translation="MTEKNPPSYKAKSGEPDLIEIMGSLRWLIALLAVLPVLLAI
VLIITKPYEARVSTLPPSLDSIAGNLRSENGLPFASGANGVYSIFVNLOSEETR
ROFFRDVYIPLSEEGASVGRDOLYALLMNSVTIOKPSKEDSRYTVVIRHDPQOAA
EMGEKYIGTAVKRSMEQEGVRRPQVKNRTIOQIDILRETKAKAREDRISLKEA
LRVAELKLENPLPIEGVDQQLUSSIMEGMLTMRGAKAIRAKITQDQRSDUFFIP
ALRGLOEKYALLLEGLTLNLDKAVAFRDQGTIEVPDPRPRKGIIVLILSLMAGFLGV
AVALAILMVRGGFRRTVDA"
2364..3632
/feature="ORF 5; similar to UDP-glucose/GDP-mannose
dehydrogenase"
/codon_start=1
/transl_table=11
/product="wbpO"
/protein_id="AA027816.1"
/db_xref="GI:20560071"
/translation="MKDLKVAVVGLGVGLPLAVERGKRTVVGFINDINOGRIAELROG
ISTLEVDAAELKEASELSFTNLQDLQCNFVITVPTIDHKQPDLTPLVKASIS
IGKVLGKDIVIYETVPATBEDCPVLEKFSGLRFNEDFFAGSPERINFQDEK
RVSSIKKVTSGSTPEIAELVSLYRIITAGTHKASSIKVAEAAKVIENTORDINIAL
INELIIFNRMGIDTBAVLKAGTKWMPFPGVGHGICGVDPVYTHTKAGSIVGH
PEILLAGRLNDGAGVAVVSQLVKMKRIHVDPGARVLMLGLTFKENCPLDKNTKV
DIVRELAENIQVDVDFPWSAEDAMHEGIVTPTGTPSHGAYDGIILLAVAHSEFKMG
AENIRKLGKAEHVLYDLKLLDEKSDLR"
3670..4695
/feature="ORF 6; similar to NAD dependent
epimerase/dehydratase family"
/codon_start=1
/transl_table=11
/product="wbpP"
/protein_id="AA027817.1"
/db_xref="GI:20560072"
/translation="WMSRYEELRKELPAOPKWLITGVAGFIGSNLETLKLQDKVV
GLDNFATGHORNLDEVRSLVSEKQWSEKFKLOGDIRNLDDNCACGVDVYLHQAALG
SVPRINDPITSNATNIDGFLNMLIARDAKQVSTFYAASSSTYGDHPLGPKVEDTIG
KPLSPYAVTKVNELYADYFCYGTGIFGRVFNVRQDNGAAYAAVTPKWTSM
IQGDVYINGDGTSDPCYIENTQANLLAATGADARNQVNIAGVGRSTLNQLFF
ALRDLGAENGVSYHREPVYRDFREGDVRHSLADISKAALKLYAPKYDVSAGVALAMP
WYIMPLK"
4692..4889
/feature="ORF 7"
/codon_start=1
/transl_table=11
/protein_id="AA027818.1"
/db_xref="GI:20560073"
/translation="MIFTVLRLDESSVWSMGAIPAFESRLVLQVLNIRGACDSLFS
RSTFSMIVFLPFTHCFLPFMV"
4896..6320
/feature="ORF 8; similar to Polysaccharide biosynthesis
protein; potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AA027819.1"
/db_xref="GI:20560074"
/translation="MLGKSLVLYFLFKSPAILTLVGLSVFTLLSPGEGVYSLTII
VWGLNTVFLQWALGVGRYLPEDSQARALIGTARALISFLVSLVIFVTFLLW
REEIGSLYVMVGLCLAQAWHDLNKLQNALLOPLTYGKMLLIKGAGSFFIGLV
YFGVGVDGLLGLTSLVSLATIPFQDNRGVSALVDKEQLTRLPAYGAPLTFLFA
FVNASDRFFIGAPLDAAGVYSVLDQYSGVTASVVHAAAPLVNKKLSKGL
POTQDLKRTFIPFVAVSPAACGLAWAPESIGSEGEFREGALKIIFLISLAF
GALKSPYDYSFQIASVAVVAVVAVVFNLIIPFEGIVGAASVSNVAFS
AIIISIFLGRVFPMPALPGDKAMKIALSVLLMAVSVASPSLESAPFLVVKVILGG
VILAAMIALDVSGMRTFLSKLIR"
6517..7488

```

```

/note="ORF_9; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AAM27820.1"
/db_xref="GI:20560075"
/translation="MKALLLEGLTRGDMHMSALVSAVPPARRDRPLWLLRELLHR
LPVIGSYGVVLLRELLSTIPTLEFLTKAPRIIDVDDAIWLHRRGIAANSIARRV
DHTVCNQYLDVFGQFGRFTIIPGTVDLRELPRERENRIVGWSGTGGYRELY
DIPLSHLFRDPEHKLRIVSDRPPFVKVPAEKVEFIRWTNETVNEIAGMDIGIM
PLADLWSRKCSYKMLLYMACGLPVVSEYGNRVDVLARGFIFYGAVDDEGWESLA
ALVKDPEARVRAQNGRDIIRHYSLDVCDLWAMVIGSVAPKGTG"
7493..9376
/note="ORF_10; similar to Asparagine synthase"
/codon_start=1
/transl_table=11
/protein_id="AAM27821.1"
/db_xref="GI:20560076"
/translation="MGIAGFWNITGLLDGNARVARQWAAAIHHRGPDESIGIWEAP
RAPHILHARLAVLELSPAGSQPHSDGCRYVLIYNGEIIYNHLARLARLSEAGVTHSWR
GSDTETLLACFAQWGVESLTKLTGVGFALALWDROKTTITLARDMGKPLVWGNQ
GVLFASLKLAKHEPLFRGDIIDRLDALFLRYGVPAPYIYKIGKLRAGSYLVLS
ERSLNETCPAAVSNAAITEALSNPFOGTDAEAVDLESOLRTSISQMWSDVPLG
APLSGVSSTVALMQOQSSRPITFSIGFDPGDEAVYAKAVAEHGTDBTELYV
NSKDALDVTSPKTICEFPDSSQIPTLIVGLARQQVTVLSUGDGGDELFGYNPY
QPTPRVWRMLERPFHSMRPPASAPADLPPELKGKRDVFASTRABEELFYRLNSHR
NTEYPVIGAQCHTALLDTPERWPRVDSFQHMWAMDVQGYMPDDILVKVDRAMANSL
ETRVPLIDHRVPELAWMLPMKIRNGKGLLREVLRYHRSRELIYRKPCKGFSVPVS
DWLRGLPKWAESLLDERRLQOEGYLDRLIRIWNHDLAGRDRHSRLWSVLMFQAW
LES"
9373..10524
/note="ORF_11; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AAM27822.1"
/db_xref="GI:20560077"
/translation="MTIVKVIHVIAGLKKVGAELMKRLIETQNGCSPEHSIISLS
DUGFQOGLIEAGISVILGVTSMRDMRVLRLIWIIFRERPDIQVTWMSDDLGG
LAARMAGIGIINGVRTDLOEGGKTTVLVVKVCWLSGFLPKYIVCAEAESRSHI
AVGNASRLVIVNGFDTLRLOATQDRQASRGAGIEASDVIIGSLGRFHPVOKHAS
FVAAGLAPRYSLKPLVGLRELLSNAELQRLIATGVAERFILLGRDVPVASCCLK
AMDIFCLHRTGETFPNVLGEAMAGLPCITDVGDAAYLLGNDGVVVPALDPNALGKG
IEDLLADVEGRALGEARQRIYSNFTWASASQRFMSLRDILVKKKA"
10568..11698
/note="ORF_12; similar to Glycosyl transferases group 1"
/codon_start=1
/transl_table=11
/protein_id="AAM27823.1"
/db_xref="GI:20560078"
/translation="MKPVLIMVNDPAFFMSHRLPVAVGAQQAQGFQVHIATRPDAVK
KIVSQFLHHEPLRSRGNPFSELYLLTYVWRLWRLRPDLVHLVTIKPIYGGIAA
RLAPGVVVAASGLGVFWMAKGLKACAFACVAMLYRRALGKKLRVIFQNPDDRDA
LIGLGAITPEKSVLIRGSGVDLTQYQAPESPETPVVTLAARLLRDKGVLEFVEANI
LQORSVPAHFQLDGDDLPNGPTSIPESELERWSECTIECLGYRODIASVFAARSLV
LPSYRGLPKLVVEASACRAVVTVDVPGCRDAIQADVTGLLVPVVRDSALADAIQRL
IISPELRKMGKAGRALAERDPAIESIVQOHLDIYRALSQA"
11839..12657
/note="ORF_13; similar to NAD dependent
epimerase/dehydratase family"
/codon_start=1
/transl_table=11
/protein_id="AAM27824.1"
/db_xref="GI:20560079"
/translation="MLLGLGSSVWMDAESIDTVVHCAARVHVMSETADPLVEPRK
ANVQGLDLAREAVSRVRFPISSIKVNGECTEPGPVTADSPNPVDVPGVSKRE
AEQALLDAETCLEVVIIRPVLYGPGVKANVQTMWRLKRGVPLPLGAIHNRSLV
SLDNLVDLITCIEHPAAQGVFLVSDGDLSTTELLRRMRGALGAPARLLVPPASWI
GAAKVLNRQAFARLUCGLSQVDIMKTRQVLGWTTPVPGVQDALEKTRFSLRDQ"
12654..13694
/note="ORF_14; similar to Glycosyl transferase group 4;
potential multiple membrane spanning domains."
/codon_start=1
/transl_table=11
/protein_id="AAM27825.1"

```

```

Query Match          4.8%; Score 57; DB 1; Length 14734;
Best Local Similarity 49.8%; Pred. No. 3.7e-06;
Matches 144; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 234 TGTGCGGCTTCTGTTAATATCAAAACCGGTACTCAATATCTTCTGTGGCGGTGGCTGCATC 293
Db 11655 TGGACGATATCTTTCGATCGCAAAATCACGCTCTGCGCAACGCTGCGCCAGCAGCTCCCATC 11596
QY 294 ATCCGGAAGCGTTTCGGTCCGGGATAAATAATCGCGAGTGCCTGCGGTCCATGCAGACACA 353
Db 11595 TTCTTCGGCACTCTCGCGACTCAATAAGACGCTGAATGGCATCCCGCAGACGACAGAA 11536
QY 354 TCCCCACGCGGTAAACGCGTCCCTGTGCATCTTCTGAATGACATCAGCGATCCCGCCCC 413
Db 11535 TCAGCGACCGGCACGAGAGCGCCCGTAAACATCAGCTTGCATCCCGGCAACCGGT 11476
QY 414 GTCTCACTGGCGAATACGCGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCAAAC 473
Db 11475 ACATCGGTGCTCACAAACCGCAGCGCCGCGACGCGCTCGACACGAGACCTTAGGAAGC 11416
QY 474 GCTTCATTTTCCGAGGCGATGACCACTGGCAATCCGGTAGACCG 522
Db 11415 CCTTCAGATAAGAGGAGAACGACTATATCGGAACGGCGGGAAGACAG 11367

RESULT 31
AF035937 16736 bp DNA linear BCT 12-JAN-2000
LOCUS Pseudomonas aeruginosa strain IAT5 O6 RpsA (rpsA) gene, partial
DEFINITION cds; Irf-Beta, Wzx (wzx), and Wzx (wzx) genes, complete cds; and
wbp gene cluster for O-antigen biosynthesis, complete sequence.
ACCESSION AF035937
VERSION AF035937.1 GI:6690121
KEYWORDS Pseudomonas aeruginosa
SOURCE Pseudomonas aeruginosa
ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 16736)
AUTHORS Belanger, M., Burrows, L.L. and Lam, J.S.
TITLE Functional analysis of genes responsible for the synthesis of the
B-band O antigen of Pseudomonas aeruginosa serotype O6
Lipopolysaccharide
JOURNAL Microbiology 145 (Pt 12), 3505-3521 (1999)
MEDLINE 20090474
PUBMED 10627048
REFERENCE 2 (bases 1 to 16736)
AUTHORS Belanger, M. and Lam, J.S.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1997) Microbiology, University of Guelph, Guelph,
Ont N1G 2W1, Canada
FEATURES
source
1..16736
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="IAT5 O6"
/serotype="O6"
/db_xref="taxon:287"
<1..805
/gene="rpsA"
<1..805
/gene="rpsA"
/codon_start=2
/transl_table=11
/product="rpsA"
/db_protein_id="AAF23994.1"
/db_xref="GI:6690127"
/translation="XPKLEEGVGLVHVSMDWTNKNIHPSKVQVQVDEVEVQVLID
ERRRISLGIKOCKSNPMDPSSQFNKGRDISGSIKSTIDFGIFGLDGGIDGLVHLS
DISWNEVEEAARVRFKKGDELETIVLSVDPERERISLGIKQLEDPPSPSYASLHEKGS
IVRGTVKEVDARGAVISGGDIEGLKAKSEISDRVEDARNVKGEEVEAKLISIDR
KSRVLSLSVKSKDVEDKDKMKELRKQEVESAGPTTIGDLIRAQMENQ"

```

```
CDS      942..1226
/codon_start=1
/transl_table=11
/product="Ihf-Beta"
/protein_id="AAF23995.1"
/db_xref="GI:6690128"
/transl_table=11
IRGFSLSLHYRAPRGNPKTGESVRLDQKGFVPHFKPGKELDRVNEPE"
1954..2811
/gene="wzz"
1954..2811
/gene="wzz"
/codon_start=1
/transl_table=11
/product="wzz"
/protein_id="AAF23996.1"
/db_xref="GI:6690129"
/transl_table=11
SDIAGNLSGSEGLPAFSANGYSIFVRNLQSEETRQFFRDVYIPSLREGASVGR
DQYALIMNSVTIQKSKESERYTVVIERHDSQAQEWGKYGIGIAVKRSMQEMQEG
VRREFQVKNRNIQQQDILRETKAKREDRISRLKEALRVAEASKLENPLIEGVDQ
QLASIMEGGLMYRGAKAIRAEIKTLQDRVSDDPFIPALRGLQEKVALLGLTLNLDK
VAVFRODGR"
3064..4332
/gene="wbpo"
3064..4332
/gene="wbpo"
/codon_start=1
/transl_table=11
/product="wbpo"
/protein_id="AAF23997.1"
/db_xref="GI:6690130"
/transl_table=11
IDSTLEVDAAELKEASELFTFNLQDKCNFVITVPTIDSHKQDPRLPLVKASE
IGVLLKGDIVIVGTVPGATEDCVPEKESGLRFNEDFPAGYSPENPGDKEH
RVSSIKKTVSGSTPEIAELVDSLYREIITAGTHKASSIKVAEAKVIENTORDNLAL
INELAIIFNRMGIDTEAVLKAAGTKWNFMFPRGLVGCHGICVDPPYLTHKAQSIGYH
PEIILAGRINDGAGTVVSQLVKMLKRIHVGDGKVLGLMGLTFKENCPLDRNTKV
DIYRELAENYIQVDVDPWVSADAMHEYGITPVGPSHGAYDGIILAVAHSEFKNKG
AENIRKLGKAEHVLYDLKYLLEDKSLRL"
4370..5395
/gene="wbpp"
4370..5395
/gene="wbpp"
/codon_start=1
/transl_table=11
/product="wbpp"
/protein_id="AAF23998.1"
/db_xref="GI:6690131"
/transl_table=11
GLDNFATGHQNRNDEVRSLASEKQWNFKEIQGDIRNLDDCNNAAGVDYVLRQAALG
SVPRSINDPITSNATNIDGFLNMLIAARDAKQVSTYAASSSTYGDHPLPKVEDTIG
KPLSPYATKYVNELYADVSRVCVGFSTIGLRYFNVRGRQDPNGAYAAVI PKWTSSM
IQGDDVYINGDSTSDPCFYIENTVOANLLAATAGLDARNQVNYIAVGCTSLNOLFF
ALRDGLAENGVSHTREPVYDFREGDVHRHSLADISKAALKGLYAPKYDVSAGVALAP
WYIMFLK"
5392..5589
/gene="wbpp"
5392..5589
/gene="wbpp"
/codon_start=1
/transl_table=11
/product="wbpp"
/protein_id="AAF23999.1"
/db_xref="GI:6690132"
/transl_table=11
RSTFSMIVLFFTHCFLFMV"
5596..7020
/gene="wzz"
5596..7020
/gene="wzz"
/codon_start=1
/transl_table=11

gene      7142..8188
/gene="wbpr"
7142..8188
/gene="wbpr"
/codon_start=1
/transl_table=11
/product="wbpr"
/protein_id="AAF24001.1"
/db_xref="GI:6690134"
/transl_table=11
ALVSAVPPARRLDRPLWLAELRLHRLPOVIGSYGSDVILQRELLSTIPTLEFLTAP
RILDVDDAILHERGIAANSIARRVDHIVCGNOYLADYFGQFGRPTTIIPGVDDTLAP
LPRRERRENVIGWSGTSGYRFLYDIEIPLSHLFRDHPWKLRIVSDRPPBEKVIAPA
EKVEFIRWTETNEVTIAGMDIGIMPLADDLWSRKCSYKMLLYMACGLPVPVSEYGM
NRDVLARGFIGYGAVDDEGRYESLAALVKDPEARVRAGQNGRDIIERHYSLDVVVDLW
AMVIGSAVFGKTR"
8193..10076
/gene="wbps"
8193..10076
/gene="wbps"
/codon_start=1
/transl_table=11
/product="wbps"
/protein_id="AAF24002.1"
/db_xref="GI:6690135"
/transl_table=11
RAPILVHAELAVLELSPAGSQWNSDCGRVLYNGEINYHNLARLRSBAGVTHSWR
GSGDSTLTACFAQGWGVESTKLITVGMFALALWDQEKTTTLARDMRGKPLYWGWN
GVLFFASLKLKALKEHPLFRGDIIDRDALALFLRYGYPPAPYSIYKIGIKLRAGSYLVS
ERSLNETCPAAVMSANAALIEALSNNPFOGTDAEAVDLESQRTSISDOMVSDVPLG
AFLSGVDSTTVVLMQQOSSRPITFSIGDFPGYDEAVYAKAAEHISTDHTELVY
NSKADUDVIPSLPKYICBFPGDSSTQPTLIVSGLARQQVTVALSGDGDSDLFGGINPY
QFTPRVYMLERFPHSMRPFASAFQDLPLPEKLGKRDVFSRTAEELFYRLNSHW
NHEYVPVIGAGHTALLDTPERWPRVDSFQHMAMMDVQGMPPDDIILKVDRAAMANSI
ETRPLIDHRFELAWRMLPKIRNGKMWLLREVLXRVHVSRELIERPKKGFSPVYS
DMLRGLKWAESLLDERRIQQEGYLDLSRLIRIWNHDLAGRDRHSRLVSLWMFQAW
LS"
10073..11224
/gene="wbpt"
10073..11224
/gene="wbpt"
/codon_start=1
/transl_table=11
/product="wbpt"
/protein_id="AAF23993.1"
/db_xref="GI:6690126"
/transl_table=11
DIFGFGGLIEAGISVDVLGTMDSMRDMPVLLRLIWRERRDPIVQTMVSHDLSLG
LAARMAGIGIINGVRTTLQEGGKSTTVLVRKVCANLWSGFLPKYIVCAEABRRSHI
AVYNAASRLMIVIPNGFDLRLQATDQRSAIRSGAGIEASDIVIGSLGRFHPVADHAS
FYAAGALAPRYSRLKVLRELLSNAELQRLIETATGYAERFILLGERQVASCILK
AMIDFICLHSRTGEPNVLGEANMAGLPCITTDVGDAAVILIGNDGVVVPALDPNALGK
IDBLIALDVEGRALCEAAQRITYSNFTMASASQRFMSLVRDILVKRKA"
11268..12398
/gene="wbpu"
11268..12398
/gene="wbpu"
/codon_start=1
/transl_table=11
/product="wbpu"
/protein_id="AAF23992.1"
```

```

/db xref="GI:6690125"
/translation="MKPVLIMI VNDPAFFMSHRLPVAVGAQQQFQVHVIATRPDAVK
KIVSQGLHELPLSRGKNPSELYLLTVWRLMLRDPDLVHLVTIKPVIYGGIAA
RLAPVGVAAVSGLGFVFNAGLKKACAPACVAMLYRRALGKKLRLVIFQNPDDRDA
LIGLALTFSKSLIRGSGVDLTQYQAPESPETPVVTLAARLRKQDGLVFEVAANI
LQRSPVPAHVLGDDLPDGPNTIIESELEWRSEGTIECLGYRDIASVFAFSHIW
LPSYREGLPLVVEASACGRAVVTVDVPGCRDAIQADVTLGLLVFVRDSALADAIQRL
IESPELRKKMGAGRALAERDFAIESIVQVQHLDIYRALGSGA"
12395..13357

gene
Query Match      4.8%; Score 57; DB 1; Length 16736;
Best Local Similarity 49.8%; Pred. No. 3.Be-06;
Matches 144; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 234 TGTGCGCTTCTGTTAATATCAAAACCGGTACTCAATATCTTCTCTGGCGTGGCTGCCATC 293
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12355 TGGACGATCTTTCGATCGCAAAATCAGCTCTGCCAACGCTGCCACGAGCTCCCATC 12296

Qy 294 ATCCGGAAGCGTTCCGTCGCGGATAAAAAATCGCGCAGTGCGCCGCTGCCATGCAGACACA 353
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12295 TTCTTGCGCAACTCTGCGCACTCAATAAGACGCTGAATGCGATCGGCCAGACGACGAA 12236

Qy 354 TCCCCACGGGTAACAGCGTCCCTGTGCACATCTTCTGATGATCATCAGGATCCCGCCC 413
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12235 TCACGACCGGCACGACGCGCGGTAAACATCAGCCTGGATTGCATCCCGGCAACCCGGT 12176

Qy 414 GTCTCACTGCGGATAACGGGCAGCGCGGAGCTGAGCGTTTCAGCAGTACCATCAACAAAC 473
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12175 ACATCGTGTCTACNACCGCAGCGCCGACGCGCCTCGACGAGACTTAGGAAGC 12116

Qy 474 GCTTCATTTTCCGAAGCGATGACCACACACTGCGCAATCCGGTAGACCG 522
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 12115 CTTTCAGATAAGAGGACAGACGACTATATGCGAAGCGGCGAAGACAG 12067
```

```

RESULT 32
AB017286/c
LOCUS
DEFINITION
AB017286 Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough plasmid
AB017286 pDV, complete sequence.
AB017286
AB017286.1 GI:46447666
.
Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough
Desulfovibrio vulgaris subsp. vulgaris str. Hildenborough
Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
Desulfovibrionaceae; Desulfovibrio.
1 (bases 1 to 202301)
Heidelberg, J.F., Seshadri, R., Haveman, S.A., Hemme, C.L.,
Paulsen, I.T., Kolonay, J.F., Eisen, J.A., Ward, N., Methe, B.,
Brinkac, L.M., Daugherty, S.C., DeBoy, R.T., Dodson, R.J., Durkin, A.S.,
Madupu, R., Nelson, W.C., Sullivan, S.A., Fouts, D.E., Haft, D.H.,
Selengut, J., Peterson, J.D., Davidsen, T.M., Zafar, N., Zhou, L.,
Radune, D., Dimitrov, G., Hance, M., Tran, K., Khouri, H.M., Gill, J.,
Uterback, T.R., Feldblyum, T.V., Wall, J.D., Voordouw, G. and
Fraser, C.M.
The genome sequence of the anaerobic, sulfate-reducing bacterium
Desulfovibrio vulgaris Hildenborough
Nat. Biotechnol. (2004) In press
15077118
2 (bases 1 to 202301)
Heidelberg, J.F., Seshadri, R., Haveman, S.A., Hemme, C.L.,
Paulsen, I.T., Kolonay, J.F., Eisen, J.A., Ward, N., Methe, B.,
Brinkac, L.M., Daugherty, S.C., DeBoy, R.T., Dodson, R.J., Durkin, A.S.,
Madupu, R., Nelson, W.C., Sullivan, S.A., Fouts, D.E., Haft, D.H.,
Selengut, J., Peterson, J.D., Davidsen, T.M., Zafar, N., Zhou, L.,
Radune, D., Dimitrov, G., Hance, M., Tran, K., Khouri, H.M., Gill, J.,
Uterback, T.R., Feldblyum, T.V., Wall, J.D., Voordouw, G. and
Fraser, C.M.
Direct Submission
Submitted (18-MAR-2004) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
Location/Qualifiers
```

```

source
1..202301
/organism="Desulfovibrio vulgaris subsp. vulgaris str.
Hildenborough"
/mol_type="genomic DNA"
/strain="Hildenborough"
/sub_species="vulgaris"
/db_xref="taxon:882"
/plasmid="pDV"
complement(699..872)
/locus_tag="DVUA0001"
complement(699..872)
/locus_tag="DVUA0001"
/notes="identified by Glimmer2; putative"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAS94447.1"
/db_xref="GI:46447781"
/translation="MFRAPFCGKDVGNWRSRYASRLIKNSVHQHFFRRIRLRFALN
DVLEGCPCACNVL"
871..1689
/locus_tag="DVUA0002"
871..1689
/locus_tag="DVUA0002"
/notes="identified by match to protein family HMM PF00991"
/codon_start=1
/transl_table=11
/product="Para family protein"
/protein_id="AAS94446.1"
/db_xref="GI:46447780"
/translation="MGAIIAIVNNKGVGKTTITNLAHALANLQOEVLVIDASQCN
TSSPPHGDVERVPAPNLYELLEDDAADVDCIHAAPESYRLAVLTHPDLACLERSVI
IQRPTGIRLMRDKLRDYAKTKYDFLIDCPNLCTFVMMVAADFVLVPLESSKY
SIDGIARTVGLIDDIRSECNKDLTLRLFLNKRPTIVARETHRLRKRPPGRVFE
TVLSASTDIQQSEYLSSETVIRSKPNSQLARLFRQLANEVVSLKEITTMQDAPSA"
1740..5135
/locus_tag="DVUA0003"
1740..5135
/locus_tag="DVUA0003"
/notes="identified by Glimmer2; putative"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AAS94445.1"
/db_xref="GI:46447779"
/translation="MPPEATANASPAQOQPKRRRVTLADRLRQASKQFESPEQIAAR
EALKMNRESALATPGDTSHGVSAAVTPDKPTTHQKVAESAADAPSPPTVTSFSP
ATEDDLGHDTSATASARASTGSDATCAASGTTPRDDAGHPLTGAVTRDVTPTETRV
THDTPVTTKATGDCDRHPATTAVTSDDTSETVATDTSDAKRHPVSPAATDASNS
HVTADVTDPDTPDSNAHRTTMAVTDGDSVTPDGELPASSASTREPKTVSSAADRAYT
APHARGROGDDTGHPMTDTVTSDGTSPPPLFVPAHGLTHGAVATDITAGAAIVE
EGPHAGVPMQROAEASAPHAAREAPTAYAGTEAGSHGHRSTAGATTABETPC
TCEAPTPTAATLHRSVARNAMGTEKVTERTVPTDPTGHPVTAATVTRHDEGL
PGAEPTPGSAESPHVTPTVPTVPTDQATPEASGTLPGATSDLTGGLTG
GSPVTAAGAEADAGTETGAPYAAAQSDATPGDDIIAPSHAARQAAGISDGAAT
VSDVTPLSAHDVTMSSEFTDDSHGHPMDNVTTTEPTVHOAHGMAPVUNSXPKEGVV
HVRSELPAHDLAHGHSQAQPTQGNVTDDASRGVSAEVTDDMSATAPVTPTVPI
PRLDTPGLAGLTSRLNSTLQVNIYNMVLEAGGAVFTGQQAQKAFDPAVRSVR
RTDLMETKGLISRESGAAMTLRLMDPLSLVMGTHASGLGAGVIGADHAGVHGI
MGCHAPYRTGRGPDMPFTPGSPAAHTPEFVAGHGTGAECRGYPAAGMPQHPHAPS
PPRDSYRTPGSPFMHAAHPVQTSPLNDQQGAAANGPDMVAPHRPAAAHQMHGGMVPR
HQPGATQQAASVSGAGTPTGACDAGFCVPSGPGCADMAFFRERYPHLARAGFGPAQ
YKQVREALSRDLDDLVESLGYAEWLEHGHGTMVDKQPIAKPADWTFKSLAMHGAC
YRRPAGYVDFLEEARIAERIAAQTAAHREQLARUREEOEKLEMALIDAEVERLAAA
PDSAEVDELLQGLPAVLRAKGLSALVQREMRRLRERLFPPEG"
5490..5843
/locus_tag="DVUA0004"
5490..5843
/locus_tag="DVUA0004"
/notes="identified by similarity to GP:14028719; match to
protein family HMM PF00216"
```



```
/product="glycosyl transferases group 1-like protein"
/protein_id="AA017419.1"
/db_xref="GI:27502139"
/translation="MTTVKVIHVTAGLKKVGGALMKRLIETQNGSBEFHSIISLS
DLGEFGGLTEAGISVDVGLWTSMDPVRLLRLIGIFRERRPDIVOTWMYHSDLLGG
LAARMAGIGGIWGVRTDLOEGGKSTVLVRKVCAMLGSLPKYIVCAEAESRRSHI
AVGNVSRMLVINGPDLTDLQDEORSAIRSGSGIEASDVIIGSLGRPHQVADHAK
FVAAGLLARYSRKLKLLVGRLLSNALQRLTEATGAERFILLGGRQDVASCUK
AMDIFCHSRSTEGPFPNGSAMGLPCTITDVGDAAYDLGNDGVVVPALDPNALGKG
IEDIALDVEGRALGEARQRIYSNFTMASQRFSMLYRDILVKKKA"
CDS
10569..11699
/notes="Orf12"
/codon_start=1
/transl_table=11
/product="glycosyl transferases group 1-like protein"
/protein_id="AA017420.1"
/db_xref="GI:27502140"
/translation="MKPVLLMIVNDPAFFMSHRLPVAVGAQAQGFQVHIATRGEAVK
KIVSQPLHELPLSRSGKNPFSBELYLLTVYWRLLWRLRDPDLHLVTKPIYVGGIAA
RLAPKGVAAVSGIGFVFAKGLKACAFACVAMLYRRALGKGGKGVIFQNPDDRDA
LIGLGAITFEKSVLRSGVDLTQYQAPESPETPVTTLAARLLRDGVLGFVEAANI
LRQSVPAHQVLQDLPNGPNTSIEPSELERPSEGTIECLGYRQDIASVFARSHIVV
LPSYREGLPKVLVVASACGRAVTTDVGCRDAIQADVTGLLVPRVDSALADAIOQL
IESPELRKKMGAGRALAERDAIESIVQHQHLDIYRALGSGA"
CDS
11840..12658
/notes="Orf13"
/codon_start=1
/transl_table=11
/product="NAD dependent epimerase/dehydratase-like
protein"
/protein_id="AA017421.1"
/db_xref="GI:27502141"
/translation="MLLGEIGSSVYVMDAESAITDVVHCAARVHMVSTASDPLVFRPK
ANVQGTDLAREAVSRGRRFIFISSIKVNGEGTEPGRPYTADSPNPDPYGVSKRE
AEQALLDLAETGLEVVIIRPVLYGVGVKATVMKLRGVPLPLGAIHNRSLV
SLDNLVDLIITCTIEHPAAVGOVLVSDGEDLSTTELLRMRGALGAPARLLPVPASHI
GAAKVLNRQAPARRCLGSLQVDIMKTRQVLGVTPPVGVQDALEKTAERSFLDRQ"
CDS
12655..13695
/notes="Orf14; potential multiple membrane spanning
domains"
/codon_start=1
/transl_table=11
/product="glycosyl transferase group 4-like protein"
Query Match 4.7%; Score 55.4; DB 1; Length 14738;
Best Local Similarity 49.5%; Pred. No. 1.2e-05;
Matches 143; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
Qy 234 TGTCCGCTCTGTTAATATCAACCGGTACTCATATCTTCTTGGCGCTGGCGCATC 293
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
11656 TGGACGATATCTTCGATCGCAAAATCAAGCTCTGCCAACGCTGCCCGACGATC 11597
Qy 294 ATCCGGAAGCGTTCGCGTGGGATAAAAAATCGCGCAGTCGCGCGGTCCATGCAGACACA 353
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
11596 TTCTTGGCAACTCTGGCGACTCAATAGACGCTGAATGGCATCCGCAGAGCAGAGAA 11537
Qy 354 TCCCCCAAGGTAACAGCGTCCCTGTGCATCTTCTTCAATGACATCAGGATCCCGCCC 413
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
11536 TCACGACCGGCACGAGCAGCGCCGTAACATCAGCTTGGATTGATCCCGCAACCCGGT 11477
Qy 414 GTCTCACTGGCGATAACGGGCACCGCGGAGCTACGCTTCAGCGATCATTACCAAC 473
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
11476 ACATCGGTCTGTCAACACGCGACGCGCGCGCGCTCGACGAGCCTTGGGAAGC 11417
Qy 474 GCTTCATTTTCGAAGGCGATGACCAACACACTGCGCAATCCGGTAGACCG 522
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
11416 CCTTCAGTAAAGAAGCGAGACGACTATATGCGAAGCGGCGAAGACAG 11368
RESULT 34
AP003599/c
LOCUS
DEFINITION Nostoc sp. PCC 7120 DNA, complete genome, section 19/19.
```

```
AP003599 BA000019
AP003599.1 GI:17134317
Nostoc sp. PCC 7120
Nostoc sp. PCC 7120
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
1
Kaneko,T., Nakamura,Y., Wolk,C.P., Kuritz,T., Sasamoto,S.,
Watanabe,A., Iriguchi,M., Ishikawa,A., Kawashima,K., Kimura,T.,
Kishida,Y., Kohara,M., Matsumoto,M., Matsuno,A., Muraki,A.,
Nakazaki,N., Shimo,S., Sugimoto,M., Takazawa,M., Yamada,M.,
Yasuda,M. and Tabata,S.
Complete genomic sequence of the filamentous nitrogen-fixing
Cyanobacterium Anabaena sp. strain PCC 7120
21595285
PUBMED 11759840
REFERENCE 2 (bases 1 to 237221)
AUTHORS Kaneko,T.
TITLE Direct Submission
JOURNAL Submitted (02-MAY-2001) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research; Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
[E-mail:kaneko@kazusa.or.jp,
URL:http://www.kazusa.or.jp/cyanobase/,
Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
FEATURES
Location/Qualifiers
1..237221
/organism="Nostoc sp. PCC 7120"
/mol_type="genomic DNA"
/db_xref="taxon:103690"
/notes="synonym:Anabaena sp. PCC7120"
complement(72..734)
/genes="all5179"
complement(72..734)
/genes="all5179"
/notes="ORF ID:all5179"
/codon_start=1
/transl_table=11
/product="O-methyltransferase"
/protein_id="BAB76878.1"
/db_xref="GI:17134318"
/translation="MTTRILGITPNLYDYLSSLSREPEILAQLRQETALQPMGRMQI
APEQGFALLVQLGAKKTLGVFTGYSLLVALPAGKLVACDINEEFTAIQAQ
RYWQAGVDHKIDLYLAPALETKLLVAGEAETDFPAFIDAKSNVDNYEERSLQI
RSGGVTAIDNVLMSGKVADPEIQDNRKTQIRAFNHKLLQDQRTLSLIPIDGGLTVLR
KN"
867..1568
/genes="alr5180"
867..1568
/genes="alr5180"
/notes="ORF ID:alr5180"
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76879.1"
/db_xref="GI:17134319"
/translation="MTLEYCATPRDIPSYISDAARYLRIPAGTIRSWTVGRHYPIS
NGSNFRPLIQICNLKPRLLSFTNLVEVHVRIRKHQIDLGKVRDALDFIDRQFOI
SPLASERFLITDGVDLPIERYGSLINASKSVKTELKDAFNAHLERIEPDDTGLAIKLY
PFRSHEENDPRFVVDPRIAFCRLVIVGTGISTRVLAERYQAGESIDELAYDYDCDR
LMIEEAIKCELPTAA"
1697..1996
/genes="alr5181"
1697..1996
/genes="alr5181"
/notes="ORF ID:alr5181"
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76880.1"
/db_xref="GI:17134320"
/translation="MEWLFVGRGWVLTGKGIKSNRLERIAVARAQIKMFTFASQ
```

```
gene
2453. .3310
/translation="SLSEEMAGILLQAIIVPMKRFVSKHPAPPIAKIYRDGHLDWMDKAQMLLEELQEF"
/genes="alr5182"
CDS
2453. .3310
/genes="alr5182"
/notes="ORF ID: alr5182"
/codon_start=1
/transl_table=11
/product="oxidoreductase"
/protein_id="BAB76881.1"
/db_xref="GI:17134321"
/translation="MAEQTKLPQPQQQPPGTESKMPKQADDAQVRGSGKUKDKVA
LITGSGIGRAVIAIYAKEGADVAVFYLSHDDAEETKNLVEEQRRRAVSIAGDITD
EAPCQAIQTVDFGKLDILINNAEQHPQSEITDITKEQLERTFRNIFSMFYLTG
AAIKHLKGSAINITSVTAYKSPOLLDSYTKGAIVAFTRSLSONLSKGRVNAV
APGPIWPLIPSPFAEKVETFGKQVPMQAGQPEEVAPSYVFLASDSSYMSGQVLH
PNGEVVNG"
3552. .3809
/genes="asr5183"
3552. .3809
/genes="asr5183"
/notes="ORF ID: asr5183"
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76882.1"
/db_xref="GI:17134322"
/translation="MSLKPLOCTELVDCARANAKOGIETAAQCCGCGDDLNRFAGEL
RKACEMNLQVKNELNLTQDMLNLGTGEIITAPDTASEL"
complement(3821. .5221)
/genes="all5184"
complement(3821. .5221)
/genes="all5184"
/notes="ORF ID: all5184"
/codon_start=1
/transl_table=11
/product="Mg2+ transport protein"
/protein_id="BAB76883.1"
/db_xref="GI:17134323"
/translation="MTETNNLSTLODVSPRELRDLVRTOMRMLLEAGDLQGAAILV
PVQADIAEIEGLPETMHALAPRLSKDBAIEVVEYLDYSVOERLIEELKSEVRDI
VDQMSDRLDELPAKVNNHLLQSLSPTEQATOLMIGVEADTAGRIMTLIEFIAL
KENNTIQUALERIRSLANASEMIYYLYVVDQARLTCIVSRELVT SQEPTIGEVMT
RDVIFVNTDTHQBEVARLIQRYDPLAVPVVDQRLAGIVTVDDVIDILEEETTKDIY
ALGGVQSSGNDPOMDLWEVAKRVLMLFVLLITNTVTGTIIKSOEDILTKVVTLTA
FIPLLTGTGNGVCAQSTVVIKGNWTDIEISLGPLOVIGREALAGALGMLGSIATI
WAYFLQRIEIVATAGSLIAISVLASISGSALPFLFRYLRLDPLMSPFITTAADV
LGVLIYFNLARVILRL"
complement(5361. .6890)
/genes="all5185"
complement(5361. .6890)
/genes="all5185"
/notes="ORF ID: all5185"
/codon_start=1
/transl_table=11
/product="mercuric reductase"
/protein_id="BAB76884.1"
/db_xref="GI:17134324"
/translation="MNSLDRVIVRVDYENQKLVAYVHPDWNVPQADNVDLVI
GAGTAGVLIVAGAGLGLGKLVALIEKHLMGDCLNVGCVPSKTIIRSRVVGIEWNG
KNLGVNIPSGIDDFPAVMARMRVRAGISHNDSAEERFASLDVDFVFLGSRFASNTV
EVAGKTLKPKKAVATGARATKPAIIGIEQAGLTWNETVPSLIRPEKLAIVGGGPIG
CELAQAPFKGSVWLTHSGSHVNLKEDNDAAQIVQOTLIKEGIRLVNAKVEVTVV
TEKRLYFSTNGHDSVTYVDEILVAGRS PNVEGLNLEAVGVYDKRRGVENDLYOT
TNPKIYAAGSDICMDWETHAADAARIIVIKNTLIFSPFGLGRSLSLVWPVYTYDPE
VAHVGLYQSDVETIKIPFSSVDRAIDQGEDGFLKIHKKGSDIEIVGATTVASHAGE
MISEVTTAIYNKIGLNKLSNVIHPTYQBAEIKKAAADTVRRITLLTPRTKLLGLFLTKF
S"
7074. .7826
/genes="alr5186"
7074. .7826
/genes="alr5186"
/notes="ORF ID: alr5186"
gene
7074. .7826
/genes="alr5186"
/notes="ORF ID: alr5186"
CDS
7074. .7826
/genes="alr5186"
/notes="ORF ID: alr5186"
hypothetical protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76885.1"
/db_xref="GI:17134325"
/translation="MIPKEQTFNRKRRLLLLGLTTLIILIVARQNIQALLQTLILM
VQSLGFGPIAYMIYINLATLFI PGSIILTKSGCLFGVFWGSSVYLIATVGAIALF
FICRLSRDWVRQIDKYPKFMIDQSVAKEGWKIVLTRLSPVPFVFLINLTAFGVTC
ISLKDYLGLGIIPGTIMVYIGSLAGDLALAGTNHQFVTPETQIWQIMQGLGLMA
TVGVTVYITKIAQKALSQKVVTETGIIKSQDAE"
7859. .8011
/genes="asr5187"
7859. .8011
/genes="asr5187"
/notes="ORF ID: asr5187"
unknown protein"
/codon_start=1
/transl_table=11
/protein_id="BAB76886.1"
/db_xref="GI:17134326"
/translation="MLTTTDKIKHPFLGVNHEEKVFEQTLTLLQLSLSLPEKVASD
ESHSHS"
7989. .8657
/genes="alr5188"
7989. .8657
/genes="alr5188"
/notes="ORF ID: alr5188"
/codon_start=1
/transl_table=11
/product="two-component response regulator"
/protein_id="BAB76887.1"
/db_xref="GI:17134327"
/translation="MNRILIAEDPRIAIEKGLRSQGTTFATVATDAYSATNMALSS
GFDMLTLGLPKQGLDVLSELRGGENIPIVLLTARDIIOQKVAGFEAGADDYLTG
PRFPELLVRVARLRQSGSQMSETVIKSGNIVULDRSKVKIQDITIELPARFT
LAETPRHPQVLSREQLDRVWGDYDIPGSNIVDVYGVYLRKKLGNLLETVRGMGY
RLRT"
8884. .10497
/genes="alr5189"
8884. .10497
/genes="alr5189"
/notes="ORF ID: alr5189"
/codon_start=1
/transl_table=11
/product="two-component sensor histidine kinase"
/protein_id="BAB76888.1"
/db_xref="GI:17134328"
/translation="MGQAKKSKSVILFHPDTQKSVKNQONTIAGTFLKRRGPFWS
TRILSWYGIILFFILFVPIPIQALYARNQVRNEBAIEKIEIFQLGLGTSNIPK
HOLDSEIRAIKQTDNRLKLPPTTKQLREFDFALGNQLPEDDTFLIALMDKPF
YKSSPRAPKEMDRDAKLRDWAKLTQPKQEEVVI PMSOVDSIVYLARPVIEIQEIMG
VIVIAHTTAGERGEVLETLAVIVQVSTFVLIPALVLAWLASGKILAPLRLITQTARKI
Query Match 4.6%; Score 54.8; DB 1; Length 237221;
Best Local Similarity 50.0%; Pred. No. 3.6e-05;
Matches 137; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
Qy 275 TCTGGCGCTGGTCCCATCATCCGGAAGCGTTCCGTCGGGATAAAAATCGCGCAGTGC 334
Db 74530 TCGGCGCGCTCTTCCCATCTGTCTAAATCTTTGGGATGTGAATAACACCAITTCGAGTTT 74471
Qy 335 GCCGGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCTGTGCACATCTTCTTGAAT 394
Db 74470 TTCTTTGAGTGCTTCTACATCTCGATGGGAATAATGAACCCATCTACTCCATCAGTAAT 74411
Qy 395 GACATCAGGATCCCGCCCGTCTCACTGGCGATAACGGGACGCGCGGAGACTGACCGTTC 454
Db 74410 AATATCAGGCGCTGCTGTATTATATAAGTTGTAATTACAGGAATTCACAAGACATGGCTTC 74351
Qy 455 AGCCAGTACCATCAACACGCTTCATTTTCCGAAGCATGACCCACACACTGCGCAATCCG 514
Db 74350 TAATAGACTAGTCCAAACCCCTTCAATTAAAGAAGGAATACTATAAACCGCTAGCAGTCC 74291
```


Qy 515 GTAGACGGTAACGCTGGGAAGCGGACCTGCC 548
Db 74290 GTAATACGTGTTTCAAGAGTAATGGGAACCTGAC 74257

RESULT 35

AE017180_24

WFCOMMENT

Sequence split into 39 fragments LOCUS AE017180 Accession AE017180

Fragment Name	Begin	End
AE017180_00	1	110000
AE017180_01	100001	210000
AE017180_02	200001	310000
AE017180_03	300001	410000
AE017180_04	400001	510000
AE017180_05	500001	610000
AE017180_06	600001	710000
AE017180_07	700001	810000
AE017180_08	800001	910000
AE017180_09	900001	1010000
AE017180_10	1000001	1110000
AE017180_11	1100001	1210000
AE017180_12	1200001	1310000
AE017180_13	1300001	1410000
AE017180_14	1400001	1510000
AE017180_15	1500001	1610000
AE017180_16	1600001	1710000
AE017180_17	1700001	1810000
AE017180_18	1800001	1910000
AE017180_19	1900001	2010000
AE017180_20	2000001	2110000
AE017180_21	2100001	2210000
AE017180_22	2200001	2310000
AE017180_23	2300001	2410000
AE017180_24	2400001	2510000
AE017180_25	2500001	2610000
AE017180_26	2600001	2710000
AE017180_27	2700001	2810000
AE017180_28	2800001	2910000
AE017180_29	2900001	3010000
AE017180_30	3000001	3110000
AE017180_31	3100001	3210000
AE017180_32	3200001	3310000
AE017180_33	3300001	3410000
AE017180_34	3400001	3510000
AE017180_35	3500001	3610000
AE017180_36	3600001	3710000
AE017180_37	3700001	3810000
AE017180_38	3800001	3814139

Continuation (25 of 39) of AE017180 Geobacter sulfurreducens

Query Match 4.5%; Score 53.4; DB 1; Length 110000;
Best Local Similarity 55.7%; Pred. No. 8.6e-05;
Matches 102; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 330 AGTCGGCGGTTCATGACGATCCCGCCCGTCTCACTGGCGATACCGGCACCGGACACTGAC 449
Db 68593 ATTTCGTGAGCAATTTGCGGGGGGTGCGCCACCGGGCACCTGAAATACCGTCCACACCGTCC 68652

Qy 390 TGAATGACATCAGGATCCCGCCCGTCTCACTGGCGATACCGGCACCGGACACTGAC 449
Db 68653 GTTACGATCTCGGCGAGCCCCCGGTTCGGGTTACACCGGACCGCGGCGCATG 68712

Qy 450 GCTTCAGCAGTACCATACCAACCGTTTCATTTTCCGAAGGACATGACACCACTGGCA 509
Db 68713 GCCTCCACCGCGACATACCAAGGCTTCATGAGGAGGAGGACGACGACGACGTCGCTG 68772

Qy 510 ATC 512

Db 68773 TTC 68775

RESULT 36

AE008327/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .10197

/organism="Agrobacterium tumefaciens str. C58"

/mol_type="genomic DNA"

/strain="C58"

/isolate="Cereon"

/db_xref="taxon:176299"

219. .1367

/gene="AGR_L_2541"

219. .1367

/gene="AGR_L_2541"

/notes="(Y17900) putative hexose transferase"

/codon_start=1

/transl_table=11

/product="AGR_L_2541p"

/protein_id="AAK89838.1"

/db_xref="GI:15159774"

/translation="MTSALFVSHTGKGAELFLADLVKAGPHSWRACFLSGGATAED

LAEGRPVMSLAGEKMSLRNNSFGALAGADYMAVAQWLSREAKHFDVCANSQ

KALFVCAALAKSRRLVILHDI VTDTAFATNRRLASLAPARIFARLVAVNSBETGR

AFTEAGGEADKVRIVYNGFDPKAKHLADAGMARLRLGDLGPPVLGRLSEWKG

QHVFLDALAMEGVQAVIVGGALFGQEAYEARREQASRLGDLGRVFLGFRSDVPEL

MASMDVVAHTSI VAPFPGRVWVAMWCGRPVATRGGVTEIIRDGTGLLVPPGDAS

ALAAALGTILSDPALAQLQGSGREDVSDRFSLQETCRSVSALLTEAA"

1396. .2514

/gene="AGR_L_2543"

1396. .2514

/gene="AGR_L_2543"

/notes="mannosyltransferase B - Synechocystis sp. (strain

PCC 6803)"

/codon_start=1

/transl_table=11

/product="AGR_L_2543p"

/protein_id="AAK89839.1"

/db_xref="GI:15159775"

/translation="MPAGGARMRILVNMPSYGGKASGVARTFSLERILLEOTGDD

YILRSPTREQLPESLRDSRLEMETPRIMVDPVLROAAMLPVLCRRHGIDVLFNV

DPGSPGGRRLTLVHSDLYFTKTPSQIGWRATLTDTDFIRLMMAGSNRVVCSSEATR

KDLARFYAAATKSLTSHSDTSLSDVDPDVTDTASPIAGRYILAVGNATSNKNFALLK

AFAMIKGKWPDLHIVHVGDEAEIAGMLVDPDLKDRILRLSGIDRQLAELVHAAC

LCVPSTVEGFCFLVLEAQQLGCPVCSNRSATPEIAGEGALTFTDPTDAQSLVAALERL

LENPELAGKLQAGYENRKLYSNQKAAESVARLFQTEA"

complement (2569. .3645)

/gene="AGR_L_2546"

complement (2569. .3645)

/gene="AGR_L_2546"

/notes="LPS biosynthesis R6bu related protein -

Methanobacterium thermoautotrophicum (strain Delta H)"

/codon_start=1

AE008327 10197 bp DNA linear BCT 28-MAY-2004
Agrobacterium tumefaciens str. C58 linear chromosome, section 131
of 187 of the complete sequence.
AE008327 AE007870
AE008327.1 GI:15159773

Agrobacterium tumefaciens str. C58
Agrobacterium tumefaciens str. C58
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
1 (bases 1 to 10197)
Hinkle,G., Slater,S.C. and Goodner,B.
Complete Genome Sequence of Agrobacterium tumefaciens C58
(Rhizobium radiobacter C58), the Causative Agent of Crown Gall
Disease in Plants
Unpublished

2 (bases 1 to 10197)
Hinkle,G., Slater,S.C. and Goodner,B.
Direct Submission
Submitted (14-AUG-2001) Bioinformatics, Cereon Genomics, 45 Sidney
Street, Cambridge, MA 02139, USA
Approximately 800 bp of telomeric sequence missing from the left
end of the chromosome and 200 bp missing from the right end.

Location/Qualifiers
1. .10197
/organism="Agrobacterium tumefaciens str. C58"
/mol_type="genomic DNA"
/strain="C58"
/isolate="Cereon"
/db_xref="taxon:176299"

219. .1367
/gene="AGR_L_2541"
219. .1367
/gene="AGR_L_2541"
/notes="(Y17900) putative hexose transferase"
/codon_start=1
/transl_table=11
/product="AGR_L_2541p"
/protein_id="AAK89838.1"
/db_xref="GI:15159774"

/translation="MTSALFVSHTGKGAELFLADLVKAGPHSWRACFLSGGATAED
LAEGRPVMSLAGEKMSLRNNSFGALAGADYMAVAQWLSREAKHFDVCANSQ
KALFVCAALAKSRRLVILHDI VTDTAFATNRRLASLAPARIFARLVAVNSBETGR
AFTEAGGEADKVRIVYNGFDPKAKHLADAGMARLRLGDLGPPVLGRLSEWKG
QHVFLDALAMEGVQAVIVGGALFGQEAYEARREQASRLGDLGRVFLGFRSDVPEL
MASMDVVAHTSI VAPFPGRVWVAMWCGRPVATRGGVTEIIRDGTGLLVPPGDAS
ALAAALGTILSDPALAQLQGSGREDVSDRFSLQETCRSVSALLTEAA"

1396. .2514
/gene="AGR_L_2543"
1396. .2514
/gene="AGR_L_2543"
/notes="mannosyltransferase B - Synechocystis sp. (strain
PCC 6803)"

/codon_start=1
/transl_table=11
/product="AGR_L_2543p"
/protein_id="AAK89839.1"
/db_xref="GI:15159775"

/translation="MPAGGARMRILVNMPSYGGKASGVARTFSLERILLEOTGDD
YILRSPTREQLPESLRDSRLEMETPRIMVDPVLROAAMLPVLCRRHGIDVLFNV
DPGSPGGRRLTLVHSDLYFTKTPSQIGWRATLTDTDFIRLMMAGSNRVVCSSEATR
KDLARFYAAATKSLTSHSDTSLSDVDPDVTDTASPIAGRYILAVGNATSNKNFALLK
AFAMIKGKWPDLHIVHVGDEAEIAGMLVDPDLKDRILRLSGIDRQLAELVHAAC
LCVPSTVEGFCFLVLEAQQLGCPVCSNRSATPEIAGEGALTFTDPTDAQSLVAALERL
LENPELAGKLQAGYENRKLYSNQKAAESVARLFQTEA"

complement (2569. .3645)
/gene="AGR_L_2546"
complement (2569. .3645)
/gene="AGR_L_2546"

/notes="LPS biosynthesis R6bu related protein -
Methanobacterium thermoautotrophicum (strain Delta H)"
/codon_start=1

gene
CDS

gene
CDS

```

/gene="AGR L 2546p"
/notes="hypothetical protein"
/codon_start=1
/db_xref="GI:15159776"
/translation="MPEQRDGFPIHRLGHTTPFFIGDRQOPWKKYKXIAVQLDPR
LVRLARVIDELQPDVNTHTSLSLTPLIWPMIRKRGILPLVHSLHDTSMCTNGSLPH
DGHICDSSKKCFARFYLHRCQCVDAVAGVGRDIVERHVRAGFTFHVPESTRYIV
NALSPDPAKPCPTARFAGPAPLVPFVGLRGIEAAKADLLIEALFPLSQWRLLVMA
VPVIGANLAVQOIKTAGLVPVFPVGVNANFFAGIDCLI VPLWPEAFGRVABAILRG
GRVIGANLAVQOIGPERQERLFAFGNAELAAEAEMRNPAMLTETPETTERIRQ
GVAPEVVAAEKLKIDVRSSQTP"
complement (3800..5323)
/gene="AGR L 2548"
complement (3800..5323)
/gene="AGR L 2548"
/notes="hypothetical protein"
/codon_start=1
/translation="MSINEARSEPVSSQSFQVAFVLAIFLLALNQ
GATDIERSIVTAAIYLLVPAIMFGPPKAGKGVLLAACFLGLGALI VQMLLOTSSM
PAWARNPWASTAMTQTIAPAATISLTADWGLMSAALPGFAMTGLVIFNTDER
AAQTIRFAIAGOMALFSIVQFALPDALGFTQKRFYGLSLTGLFVNRTAATFFGL
ILTLATLHLSLLAPDMVMKALVANRLTPADOKKLRKSVFVIGLWAGFSFIALML
TGSRAGLASSLAALIPLILLTVFNSAPKTRNGASRRKORGSRRRAALVIAVAIALF
ALFANVAIRMETLEDMDRFVCMYMGCIARAITDNWPLGSLGSFAFIYAPYHARGCV
DAVTAHNVYAEGLTLGNAPFFYVAFVVLQTLPIRGTRKRRNRYRASHGLGAL
LLVLLHLSDFSIQIFGMAYAVFLAPVITLCNPPGPERDARKQROSSAVVISE"
complement (5402..6130)
/gene="AGR L 2549g1"
complement (5402..6130)
/gene="AGR L 2549g1"
/notes="hypothetical protein"
/codon_start=1
/translation="MPEVPAIIIRATRIDITWTRDGVMLRKVLAVPVLVIVSVFL
SVAAEAFSQRSDIVAMAKIADNNGIAPALLAATVPKLSVVTETKICRSDIVKAG
LRULIIDLANGADSPASGMARVDFAETFIHSLFCLPANGDWLRLAMVRLRNAS
PMEVAVLMFNSQLYGPADANLIRGFVWMQKFORDALPOVAARDADTAVVCGKEGI
LRWTLAAVCVKPPSGTKRPATLP"
complement (6072..8348)
/gene="AGR L 2554"
complement (6072..8348)
/notes="AGR L 2554"
/notes=" (AF067140) exopolysaccharide polymerization protein
Psep"
/codon_start=1
/translation="MLQRHDPITMSDYDKSGLTAFBIDIDAALAIIRROWRVVLAIA
VAGIAGFAATVPIYVSATALLIDRNNISOIVEQLSTIGGVVEDEASILSOVELQS
ETIGLAVLSKLTEQBFNATRASLLSIFGIRSLNVNSQWSPFTEKGAVIDDGLT
KRSLLRLNDLSVKRIGRTYALELTYNSTSPVLAQOIVNAVASYLLDKLMSKEAT
RRASDLRIALQALDLDIAVQKFAEHLNLTGNNGLISDQDLAESALILA
QSETAKARVQRIEHLATDDVADVITDILSSVANDLRKKYLSKLEETITRLIG
SHIIOAVRLNMOEYRRLMFOEISRIAQSYKSDLEVSEAKESLAEVAKADTISNS
ASTQVOMELQREASTYKNMOTFLQRYQEAQOQSPVTEARVLSKAMPVMPSPKP
NKPMLIFWVMAAGGGLAIFREFRDFRTGDQVRDVLGLELGNTPILPNQPTT
EAQENHGPGLTRPVSARVADHPLSFAETLRDIAIDLGPATKAGRI VGVVSSL
PESGKSTISINLAQAGQCARVLLDADIRNPGATRAMAHAEGLLEVLGERSVQ
DVLRLDEKTYLAPLPTVVKQRPVHPSSELLTSAQHKLAEASAFYIIVDPLPGPV
VDAMAGRIDGFI FVTEWGTARRAVNTVENEVHRRKKCLGLVILNKVDYTEKLIVR
AYGSSFEVHSRYTVYHD"
complement (8453..9445)
/gene="AGR L 2556"
complement (8453..9445)

```

```

/gene="AGR L 2556"
/notes="hypothetical protein"
/codon_start=1
/translation="MLDNHIQHWMTKSYFKDLESMPRIHNGSMLVGMFVASLGVGF
SSVNAQTVETLPPRAAPQOTPAVESAACTAPVTADESTAAFLARVNTNPDGVSISA
AELFKQSPQOAIVKOLIVTASSTDAATADTVTKVTSGFPIIAVSTISAAQFSEQLSALI
LSDSGQPTVALAKVGNKQFSEVVGEGLALAHATAMRNGNTILILAAQLQKTKGPGV
PSDLLIAFSENIOGTAAAPAAAAPAAAPAAPIGGAGSSPANSIGSGTAGTEGTA
QAGGAAPVGGGSGVGTTSGGTSPSTTTNNFTTINNFTSGNTTTPVSTSPV"
complement (9561..10130)
/gene="AGR L 2557"
complement (9561..10130)
/gene="AGR L 2557"
/notes="hypothetical protein"
/codon_start=1
/translation="AGR L 2557p"
/protein_id="AAK89845.1"
/db_xref="GI:15159780"
/translation="MNFSGAKAGPIYQIYLYPYRLTMKNFTKAPATGAQRDLVQV
MRVAAVMIVAHADVNWEFTLPLRALRGGVDLFFVICGFVIYASRPYPAAAGG
RGAFLRRRLRIVPLVYFVLTVLMAAAVPTIIGAKAPPEWSNLVYLVIVISTVAIA
HAFVFPVETPANLWLRDLRLITSPRKAVS"
ORIGIN

Query Match      4.4%; Score 52.2; DB 1; Length 10197;
Best Local Similarity 48.2%; Pred. No. 0.00013;
Matches 178; Conservative 0; Mismatches 188; Indels 3; Gaps 1;

QY 255 AACCGGTACTCATATCTTCTCTGGCGCTGGTCCCATCATCCGGAAGCGTTCGCGTCGG 314
    |||||
Db 1312 AACCGGTCACTCACATCTCTCCGCGCTTGTGCCAGCGCTTTCGCCACAGCAGGATCA 1253
    |||||

QY 315 GATAAAAAATCGCGCAGTCGGCGTCCATGCAGACACATCCCCCAACGGGTACACGGTC 374
    |||||
Db 1252 GACAGATCGTACCAAGTCGCGCGCAGCGCGATGCCATCCCCCGGCGCAGAGAAGA 1193
    |||||

QY 375 CTGTGCACATTTCTTGAATACATCAGGGAATCCCGCCCGTCTCTACTGGCGATAACGGGC 434
    |||||
Db 1192 CCGGTTTCTCCGTACGAATGATTTTCAGTAACCCCGCCGCGCGCTGGCGACGCGGC 1133
    |||||

QY 435 AGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTT--CATTTTCCGAAGGC 491
    |||||
Db 1132 CGCCCGCACATCATCGCCTCGACAACCCCGCGCGAGAGGTTTCGGCAACGATCGAGGTA 1073
    |||||

QY 492 ATGACCAACCACTCGCAATCCGTAGACCGGTAAACGCTGGGAAAAGGGCACCTGCCATT 551
    |||||
Db 1072 TGGGCAACCACTCCATCCAGCCCATAGTTCGGGCACATCCGAACGGAACCAAGGAAG 1013
    |||||

QY 552 AACACATCTCCGCTCATCTCCAGGTTCTGTCTGTGTGACGCGAGCTGCTTCGTATTCT 611
    |||||
Db 1012 CGGACGCGGCCATCCAGCCCCAGACGCGACGCTGTTTCGGAATGCGCGCTCGTAGGCT 953
    |||||

QY 612 TCACGCCCG 620
    |||||
Db 952 TCCTGTCCG 944
    |||||

RESULT 37
AE009285
LOCUS
DEFINITION
  Agrobacterium tumefaciens str. C58 linear chromosome, section 55 of
  187 of the complete sequence.
ACCESSION
  AE009285
VERSION
  AE009285.1
KEYWORDS
  Agrobacterium tumefaciens str. C58
  Agrobacterium tumefaciens str. C58
  Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

```

```

Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
1 (bases 1 to 12006)
Wood,D.W., Setubal,J.C., Kaul,R., Monks,D., Chen,L., Wood,G.E.,
Chen,Y., Woo,L., Kitajima,J.P., Okura,V.K., Almeida Jr.,N.F.,
Zhou,Y., Bovee Sr.,D., Chapman,P., Clendenning,J., Deatherage,G.,
Gillet,W., Grant,C., Guenther,D., Kutyavin,T., Levy,R., Li,M.,
McClelland,E., Palmieri,A., Raymond,C., Rouse,G.,
Saenphimmachak,C., Wu,Z., Gordon,D., Eisen,J.A., Paulsen,I.,
Karp,P., Romero,P., Zhang,S., Yoo,H., Tao,Y., Biddle,P., Jung,M.,
Krespan,W., Perry,M., Gordon-Kamm,B., Liao,L., Kim,S., Hendrick,C.,
Zhao,Z., Dolan,M., Tingey,S.V., Tomb,J., Gordon,M.P., Olson,M.V.
and Nester,E.W.
The genome of the natural genetic engineer Agrobacterium
tumefaciens C58
Science 294 (5550), 2317-2323 (2001)
21608550
11743193
2 (bases 1 to 12006)
Wood,D.W., Setubal,J.C., Kaul,R., Monks,D., Chen,L., Wood,G.E.,
Chen,Y., Woo,L., Kitajima,J.P., Okura,V.K., Almeida Jr.,N.F.,
Zhou,Y., Bovee Sr.,D., Chapman,P., Clendenning,J., Deatherage,G.,
Gillet,W., Grant,C., Guenther,D., Kutyavin,T., Levy,R., Li,M.,
McClelland,E., Palmieri,A., Raymond,C., Rouse,G.,
Saenphimmachak,C., Wu,Z., Gordon,D., Eisen,J.A., Paulsen,I.,
Karp,P., Romero,P., Zhang,S., Yoo,H., Tao,Y., Biddle,P., Jung,M.,
Krespan,W., Perry,M., Gordon-Kamm,B., Liao,L., Kim,S., Hendrick,C.,
Zhao,Z., Dolan,M., Tingey,S.V., Tomb,J., Gordon,M.P., Olson,M.V.
and Nester,E.W.
Direct Submission
Submitted (27-SEP-2001) Department of Microbiology, University of
Washington, 1959 NE Pacific Ave, Box 357242, Seattle, WA
98195-7242, USA
Location/Qualifiers
1. 12006
/organism="Agrobacterium tumefaciens str. C58"
/mol_type="genomic DNA"
/strain="C58"
/isolate="U. Washington"
/db_xref="taxon:176299"
100. 360
/gene="Atu3552"
100. 360
/gene="Atu3552"
/notes="putative; ORF located using Glimmer"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL44364.1"
/db_xref="GI:17741959"
/translation="MTKIYFSKSTGTGGTSSDFRGAQVVRVCEKQLSRVTVFNSFK
CRREAIIGFIIVDNLIDNIYFISKWILERLQTPETSSTP"
727. 1785
/gene="Atu3553"
727. 1785
/gene="Atu3553"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL44365.1"
/db_xref="GI:17741960"
/translation="MTPDKRNSPVFORLRRIVASNIIRLTIIYDNLNPNVYVPLDRL
ADFWACDGDIDIFEGSNEPLNTKDPONAPRISAKHOAALYNVAHSNPQISVPAAP
SVYLTFRALADLDKADCYDNIHPYAGMEHPETGPGALSKSIAASAHI AAGPKVLAT
EMGYHLSQTKTLHFVYTEGIKARYMPMLLMSFISGIRSIYIELVSSFAADATNPE
SSFGLRHLDLSRTPAYEAVRGLSLCLCKANPSTISVEREAVFLDADQAQLRSMPLTRQDQ
GGQVRLSVSDQLTVLEVF"
1992. 2429
/gene="Atu3554"
1992. 2429
/gene="Atu3554"
Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL44366.1"
/db_xref="GI:17741961"
/translation="MRAVAVMIVLHADVNMEFTLPLRALRLGGVDLFFVICGFV
IVASRPYFAAAGRSAPFLRLRFLRVPLYFWFLTVRLAMAAVPTTIAGAKAPPEWSN
LVLVIVISVTAIAHAFHVFTETPANTLRLDLITSPRKVAVS"
2611. 3537
/gene="Atu3555"
2611. 3537
/gene="Atu3555"
/notes="putative; ORF located using Glimmer"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAL44367.1"
/db_xref="GI:17741962"
/translation="MPRIHNGMLVGMFVASIGVFSVSSVANAQTVETLPAAPQOTPA
VSAATGAVPTADESTAAFLARVNTNPDGVSISAELFKQSPQTAIVKDLVTIASS
TDAATADVTKVTSQFPPIIAVSTSAAQFSEQLSALILSDSGQLPTVVAAQKAGVGNKF
SEVVGELAHATANGNTTILAAAIQLQTKGPGVPSDLIIAFSENIOGTAAAGPAAA
AAAPAAASPIGGAGSSPASANSIGSGTAGTEGCTAQAGGAAPVGGGSGVGGTTSGG
GTSPTSTTNNFTTINNFTSTGTTTTPVSTSPV"
3642. 5918
/gene="Atu3556"
3642. 5918
/gene="Atu3556"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="exopolysaccharide polymerization/transport
protein"
/protein_id="AAL44368.1"
/db_xref="GI:17741963"
/translation="MLQRHDLTMSDYKSLTAERIDIDAALAIIRORVRVLAATA
VAGAIGLAFAATAVPIYSATATLIDRNNNSQIVEQLSTIGGVVEDENSLISQVEVLF
ETIGLAVVDSKLITENQEPATRASLLSIFGTRSLNVNSQWFSPTKBAVDDGTL
KRSLSDLNDLSVKRIGRTYALTYNSTSPVLAQIIVNVAASAYLLDKLNSKYEAT
RRAASDLSDRIAEQLRQALDTLAVQKFAEHNLI TTGNGLLSDQOLAESNSALILA
QSETAKARAVORIEHLATDDVDVAVTDLSDSSVANDLRKKYLESSKLEITRRIG
SSHQAVRLNMQEYRLMFQEIISIAQSYKSDLEVSEAKEKSLAESVAKATDINS
ASETQVQRELQREAEYTKMYQTFQRYQEAQQQSPVTEARVLSKAMPPTMPSK
NKPMILVLFMWGAAAGGAIAPREFRDFRFTGDOVRDLGLEFIGNTLIIPNQPT
EAQENGHPGLTRPTSVARVADHPLSSFAETLRSLAIDLGI PAKTGARIVGVSSL
PSEGKSTISINLAQLAAGQARVLLDADIRNPGATRAMARHAAEGLELVLEGRSVQ
DILLREKTYLAFPTVVKQVPHSSSELLTSQMHKLAEASAFDYIIVDLPPLGPV
VDARWAGRIDGFI FVTEWKGTRARRAVNTVENEVHKKCLGLVILNKVDTEKLKYR
AYGSSEYHSRYRYHD"
6667. 8190
/gene="Atu3557"
6667. 8190
/gene="Atu3557"
/notes="identified by sequence similarity; putative; ORF
located using Blastx/Glimmer"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAL44369.1"
/db_xref="GI:17741964"
/translation="MSLNARSSEPFVSSQPSMAFONYARSAIVSFWAIFLLALLNQ
GATDIBSRVVTAAIYLLVPAIMFPGPKAGAKVLLAACFLIGALIIVMLLQTSMS
PAWARPNPASTAAMETQIAPATISLTADDWLGLMSAALPFGAFTGIVATFNTDR
AACTLRWPAIAGWALAFSIVQALPDLALGFTQKRFYGLSLTGLFVNRVNTATFGL
ILTLTLAHLKSLADPAMVAKALVANRLTPADQKRLKRSVFIVGLAGSFIALML
ILTLTGAGIASLAALI FILLITVENSAPKTRNGASRRRKGSRRAALVIAVALF
ALFANRVAIRMETRLEDWMRFYMPGIRAITDNPGLSGLSLSEFAETIAPYHARCGV
DAVVTAHNVYAEGLLTGAAPPYVAFVLTITFIRGTRKRRRYRASHGLAAL
LVLVLSILDPSLIQIPGFAMAYAVFLAPVITLCINPPGVVERDARKRQRQSSAVISE"
8207. 9421

```

CDS	/gene="Atu3558" 8207.. .9421 /gene="Atu3558" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer" /codon_start=1 /transl_table=11 /product="glycosyltransferase" /protein_id="AAL44370.1" /db_xref="GI:17741965" /translation:"MRLLQTSLSFSDRVCGSAIFVEDLARGLADKGHVTVAVAISRE QQMPEDQRGAFAIRHGLTFFFGDMRWQPEWRRKKYTKIAVQLDPLRLVRKLARVIDEL QQDVNTHSUSEITPLIWMIRKRGIPLVHSLDFTSMCTNGSLFDHGHI CDGSSKKC RATSYLHRRCQCSVDVAGVGRI DIVERHVRAGFETHVPESRRTVIWNASIPDPAPKPC TASAPAGPVLVGYLRTEAAKGADLLTEARFLPSQGRWLWAGRAPDGIEAQO KTAGLPVERPGYEANNFRAGIDCLIVPLLWPEAFGRTVAEAILRGVPVIGANLAGVA EIQGPGEERLPAFGNAELAARMAMRNPMALTETPRETRIROGVAPVVVAAAYE KLXSDVRSSTQP"
gene	complement(9476..10570) /gene="Atu3559" complement(9476..10570) /gene="Atu3559" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer" /codon_start=1 /transl_table=11 /product="glycosyltransferase" /protein_id="AAL44371.1" /db_xref="GI:17741966" /translations:"MRLLNMPSQYGKAGVARVTFSLLELTQDDYILRSPMT REQPLESRDLSELMETPRIPFVDLRFQAAMLVLCRRHIGIDL FNVDPFGSPLG KERLLTLVDLYPKTTPEQIGWAATLTDFIPRLMAGSNRVVCVEARKOLARPY AAATKSITIHSDSTLSVDPPVDTASP IAGRYTLAVGNATSKNPNALKAKPAMICKY WPDLHVHGRDRAEEIAGMLVPDKL KDRLLIRLSGIDDRLQLAELYHAACLVCPSTYE GFCPLVLEAOQLGPCVCSNRSATPEIAGEALLTGALSGILVAALERLENPELAG KLQAGYENRKLYSWQKAESA YRLFQTEA"
gene	complement(10623..11771) /gene="Atu3560" complement(10623..11771) /gene="Atu3560" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer" /codon_start=1 /transl_table=11 /product="glycosyltransferase"
CDS	Query Match 4.4%; Score 52.2; DB 1; Length 12006; Best local Similarity 48.2%; Pred.No. 0.00013; Matches 178; Conservative 0; Mismatches 188; Indels 3; Gaps 1; Qy 255 AACGGTACTCAATATCTTCCTGGCGCTGGCTGCCATCATCGGAAGCGTTCCGGTCGG 314 Db 10678 AAACGGGTCACTACATCCTCCGCGCGCTTTGTGCCAGCCTTTTGCGCCAGAGCAGGATCA 1073 Qy 315 GATAAAAAATCGGGCAGTGGCGCGGTCCATGCAGACACATCCCCCACGGGTAAACGGTC 374 Db 10738 GACAGGATCGTACCAGTGGCGCGCGCAGCGCGATGCATCCC CGCGCGGCACGAGAAGA 1079 Qy 375 CTTGTGCATTCTTCTGAATGACATCAGGATCCCGCCCGTCTCACTGGCGATAACCGGG 434 Db 10798 CCGGTTTCTCCGTCACGAATGATTTTCAGTAACCCCGCGCGCACGCGTGGCGACGACGGCG 1085 Qy 435 ACGCCGAGACTGACGCTTCAGCCAGTGACATACCAAACGCTT---CATTTTCCGAAGGC 491 Db 10858 CGGCCGACATATCGCCTTCGAACAACCCCGCGCAAGGTTTCGCAACGATCGAGGTA 1091 Qy 492 ATGACACACCACACTGGCAATCCGCTAGACCGGTAAAGCTGGGAAAAGGGCACCTCGCAATT 551 Db 10918 TGGGCACACATCCCATCGAGCCATTAGTTCGGGCATCCGAAACGGAACCAAGGAAG 1097 Qy 552 AACACATCTCCGCTCATPTCCAGGTTGTCTGTCTGTGACGACAGCGTCTCGTATTCT 611 Db 10978 CGGACGGGCGCATTCAGCCCCCAACGACGACGCGCTGTTTCGGAATGCGCGCTCGTAGGCT 1103

TITLE	Rubres,X., Tomas,J.M. and Regue,M. Cloning and characterization of two Serratia marcescens genes involved in core lipopolysaccharide biosynthesis
JOURNAL	J. Bacteriol. 178 (19), 5741-5747 (1996)
MEDLINE	96422003
PUBMED	8824620
REFERENCE	2 (bases 1 to 20693)
AUTHORS	Coderch,N., Pique,N., Lindner,B., Abitau,N., Merino,S., Izquierdo,L., Jimenez,N., Tomas,J.M., Holst,O. and Regue,M.
TITLE	Genetic and Structural Characterization of the Core Region of the Lipopolysaccharide from Serratia marcescens N28b (Serovar O4)
JOURNAL	J. Bacteriol. 186 (4), 978-988 (2004)
PUBMED	14761992
REFERENCE	3 (bases 12629 to 15061)
AUTHORS	Regue,M.
TITLE	Direct Submission
JOURNAL	Submitted (27-MAR-1996) Microbiologia i Parasitologia Sanitarias, Facultat de Farmacia, Universitat de Barcelona, Av. Joan XXIII s/n, Barcelona 08028, Spain
REFERENCE	4 (bases 1 to 20693)
AUTHORS	Pique,N., Coderch,N., Abitau,N., Aguilar,A., Merino,S., Tomas,J. and Regue,M.
TITLE	Direct Submission
JOURNAL	Submitted (30-APR-1999) Micoribologia y Parasitologia Sanitarias, Universidad de Barcelona, Av. Joan XXIII s/n, Barcelona 08028, Spain
REMARK	Sequence update by submitter
REFERENCE	5 (bases 1 to 20693)
AUTHORS	Regue,M., Pique,N. and Abitau,N.
TITLE	Direct Submission
JOURNAL	Submitted (26-OCT-2001) Micoribologia y Parasitologia Sanitarias, Universidad de Barcelona, Av. Joan XXIII s/n, Barcelona 08028, Spain
REMARK	Sequence update by submitter
COMMENT	On Oct 26, 2001 this sequence version replaced gi:4753134.
FEATURES	Location/Qualifiers
source	1..20693
	/organism="Serratia marcescens"
	/mol_type="genomic DNA"
	/strain="N28b"
	/db_xref="taxon:615"
	<1..775
gene	/gene="gmhD"
	/note="synonym: rfaD"
CDS	<1..775
	/gene="gmhD"
	/codon_start=2
	/transl_table=11
	/product="ADP-L-glycero-D-manno-heptose epimerase"
	/protein_id="AAL23753.1"
	/db_xref="GI:16445332"
	/translation="EDFLIQIMAGEBFEIEAIFHEGACSSTTTWDGKYMNNYQYS KELLHYCLEREIPFLYASSAATYGRSDPESIESEYEQPLNVGYGSKFLFDEVYRIL PEANSOIVGPRYFNVPGRGHGKSWASVAFHNTOLNNGESPKLPEGSDGKRDVY VGVADVNLFWENGSGIFNLGTGRAESFOAVADATLAKHKGSIYEIYPPPEKLGR YQAVTQADNLNRKAGYDKPFTVAEGVTBYMAWLARDA"
gene	785..1831
	/gene="waaF"
CDS	785..1831
	/gene="waaF"
	/note="complements a Serratia enterica serovar Typhimurium waaF mutant"
	/codon_start=1
	/transl_table=11
	/product="heptosyltransferase II waaF"
	/protein_id="AAL23754.1"
	/db_xref="GI:16445333"
	/translation="MKILVVGPSWVGWMMMSQRLYTLRLARYFOALIDVNAWCRPL LARMEVNOALAMPGLGALGERRLRALRANRYDRAYVLENSFKSALVPFFADL PORTWGRNGEYGLLNDVLDKAAFLVMQRYVALAYDKGRVORADDLQPQLWLPOL RVSDERTAEFTTAPNLTDSPITVGPCGAEFGAKRWPHVHYAALAQRLTESGQIVL FGSAKHGEAGEQIRAAQLQDDARDFCLNLAGTKAQEONVILLIACRAVNSDGLMIVA AALNKPLIAYGFSPPDFTPTPLSDKARVIRLISGYHKVRKGDAEQGHQSLIDIQQQ
gene	VLDALTPLLVASEE"
	/gene="waaC"
CDS	1834..2799
	/gene="waaC"
	/note="complements a Serratia enterica serovar Typhimurium waaC mutant"
	/codon_start=1
	/transl_table=11
	/product="heptosyltransferase I waaC"
	/protein_id="AAL23755.1"
	/db_xref="GI:16445334"
	/translation="MQVLIVKTSMDGDLVLTLPALDALQAIPIDIRFDWVBERGFSOI PTHPAVDRIPIVPIRWRKRWGSGSTRQCRDFKALQERRRYDVVIDAQLMKSAL ITRVAKSGKHGPDCKSAREPFASFWFNVRHEIDKQRAVERTELFAKSLGUKPGSY ASYAAARFLSQPPVDAGVYLPKLHATTRDDKHMPQONRELIATDSGLKLPWG GHAEHQRLARLAEFGPHVEVLKSLQQVADVLGARGVYVSDTGLSHLTAAALDKPNI TLFQPTDPLGILGYGQNHSLISPEKSMATIDATTAQALQKVA"
	2796..3962
	/note="Orf4"
	/codon_start=1
	/transl_table=11
	/product="putative glycosyltransferase"
	/protein_id="AAL23756.1"
	/db_xref="GI:16445335"
	/translation="MKKLHINLKGWGVVERLFOVINDTDCGNOVLCTISGDIGREI RKQPAHQPVITFANRLINALPCQPLFRKFLKWKIERANADVIVUWDLVPLGAAP KRGLVYDHGCSWRYPKQKTLRFFAMLDGVISSASHASKRVMELRPNLPCPNHVIN RIKTPAGIDINTPKTISQPIRIGTASRLVSKGISVSLMMQELLRHGHVTLVAVGK PRAPAEALAAARLQLDGRVTFSGYQDDVAGFNRTIHYMSTPITEPFGSLCMEISLYFG VPVIFPQDGOPEAVKDGVCYGLVPSVTIEOHRQLTDIEVDFPHEVYDPLTSLVSP KLLSHIDCADAVETLVAGETTYQTLNRQRYPAEHFNVAQFKAEDDTLASFIV"
	3991..5232
gene	/gene="waaL"
CDS	3991..5232
	/gene="waaL"
	/note="similar to Escherchia coli K-12 waaL"
	/codon_start=1
	/transl_table=11
	/product="lipid A core:surface polymer ligase waaL"
	/protein_id="AAL23757.1"
	/db_xref="GI:16445336"
	/translation="MLKQKTPHAPSYLIVGCAIAFCTIPFGSATGRNLFYVSSYIA FIACLYPRYFSNVKXNLLLPALMFCVGMGTILMWHFKQGEVITIRSYMSTGKLQ LATAFILLALNERLCVQRLLIVALTGLAINGSLYQGLRDLIDFRAVELNFRATVA AYLMTADLMMQAILMLRTRYRLVLYIAAFLLSALVLTGTRAMLVYVPAICLSL LATKHLVSRKHKLAVASVPLLVACGLVFKQIERIOAFETDMLMDQPQENSIIL SLKMOTLAWRTGSOAPWGQSAEORGEIRAIVAQOPLSGVMPIYINHLHNELEY SLKGVWGLLLALYIGLFFSSFRPQRNALLLSVASLSFVYGLSDVIFPFGTEGTMFC LALIAGVLSVAKPPSAEQTA"
	5229..6209
	/note="Orf6; similar to TerB (WbcK) protein from Yersinia enterocolitica"
	/codon_start=1
	/transl_table=11
	/product="unknown"
	/protein_id="AAL23758.1"
	/db_xref="GI:16445337"
	/translation="MKTHOPTLLSVIPFFYNNEAFVIAALDSLFTQISDIEVVIDD GSTDASGALVSRVLAQORPVAFISQANGGIAHARNVQLQHAAGRYLTFLDGDLLS DLYALIRPLLIADYDYLIDENYKFTQEPKPNRATOPAAVDFSSLGSLQPLFA RSNWHLWSRYKXTLLAGETPFEDGRYEDVIFPPQYFKTRIAHLDCELYFYRDNQO GITRNIKPKDIEDLLFAMEKMLRFVAQHGPDEPLRAKMLANLMLANCESEVSKSKAVYG YHYAPATQILIRRAATVCRNSDVCPCKVRQMYARVDTFLSKLRGKKR"
	complement(6246..7340)
	/note="Orf7; LPS from a non polar orf7 mutant has reduced levels of D-D-heptose"
	/codon_start=1
	/transl_table=11
	/product="heptosyltransferase"
	/protein_id="AAL23759.1"
	/db_xref="GI:16445338"
	/translation="MSKWKIKVADWFLRYTAKYGRFQDKATFDCAQQLDNIVYST

gene	TALGDFMFTPALRAMRERYNAHITLVVHKKNRELVENGRYYDRVVWVWSKIITIRP LKLRSRQPDIALILHSHPLPYDIISAVWAGAKYIIRDNYSSGICGLEPWLWYIPIHY YGHFVRKLELISIGLRSDDPDPIVDVQVQKSSDHPETIGFOLGASRESQWVS HYARLDELVAKHQNVRIVLIGAPLDGKAQPMWELIGPQTQACVQDNQIGKGTGLQLL TLDSFHLVGTGTGHLHVAUKVITISLFTVADSRSTGPIQDPBELLRIQVSRGTGY TMPLAESPMGVIPQRIYDELVAHVPLAR"	ACCESSION VERSION KEYWORDS SOURCE ORGANISM	AE017269 AE017194 AE017269.1 GI:42736432 Bacillus cereus ATCC 10987 Bacillus cereus ATCC 10987 Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.
CDS	8516. .9598 /gene="waaQ" 8516. .9598 /gene="waaQ" /note="complements a non polar Klebsiella pneumoniae waaQ mutant; similar to WaaQ proteins from the Enterobacteriaceae" /codon_start=1 /transl_table=11 /product="putative heptosyltransferase III WaaQ" /protein_id="AAL23760.1" /db_xref="GI:16445339" /translation="MNDAPALTSAPSVQRIILIKLRHHGDMLLVTPVISSLRQNPQ ARIDVLLYQETQMLASNPELSTFLAIDROWKKQCARHLGHELALRLKAKRYDLV VNLADWRSAIITRLGARIRLGFDPKRGFLWRHCHTQLVPDVNDNHLHTVEONLS LLAPLNPALNEHVTSYDPDQWQTCQELLQKQGVAGSYIVQPTSRWFKWSEEM RAKTLQADGHQLVITSGDAREKAWERILALCPQGVISLAGQTLRLQALALDH AKLFIGVDSVPMMAALQTPCVALFPGSKLTFWRPQATGAVIMAGDFGELPDPDAI DTGTDERYLDLPTDAVIAAARSTLA" 9595. .10722 /note="similar to Klebsiella pneumoniae C3 putative glycosyltransferase and WaaG from Salmonella enterica serovar typhimurium, and Escherichia coli strains K-12, R2, R3 and R4; Orf9; LPS from double non-polar orf9-orf10 mutant lacks Gala, has reduced Glc levels and contains O-antigen" /codon_start=1 /transl_table=11 /product="putative glycosyltransferase" /protein_id="AAD28801.1" /db_xref="GI:4753135" /translation="MKAPLLAIVRRKYRPGDGAERFVSRAKALEOODLDLVNITREW QGDANPNWHILCNPLKIGISRGEFAVAARALWOKERFDLVQSHERIPGCDIYRAG DGVRWELLQRAELLPEWRKWLFSNRYHYVMCAERAMYAPELKAVICNAEMIKRE IIADFGVPADKITYIYNADNQKFPPEADSAQRRLREQIYIQQAHCLIFVSGFERK GLAAIRAATAATDSDHLLVVGKQAEKRYRALAQSLGCGDRIHFMGQKTLFPYQAAH ALLPLTYDPFPNVILEAMSCGLPVITSTTCGGAETITPQNGFVTDALDVPATEAI	REFERENCE AUTHORS TITLE JOURNAL PUBMED REFERENCE AUTHORS TITLE JOURNAL FEATURES source	1 (bases 1 to 294300) Rasko, D.A., Ravel, J., Okstad, O.A., Helgason, E., Cer, R.Z., Jiang, L., Shores, K.A., Fouts, D.E., Tourasse, N.J., Angluoli, S.V., Kolonay, J., Nelson, W.C., Kolsto, A.-B., Fraser, C.M. and Read, T.D. Direct Submission Submitted (19-FEB-2004) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA Location/Qualifiers 1. .294300 /organism="Bacillus cereus ATCC 10987" /mol_type="genomic DNA" /strain="ATCC 10987" /db_xref="ATCC:10987" /db_xref="taxon:222523" 112. .915 /gene="phnX" /locus_tag="BCE1439" 112. .915 /gene="phnX" /locus_tag="BCE1439" /EC_number="3.11.1.1" /notes="identified by match to protein family HMM PF00702; match to protein family HMM TIGR01422; match to protein family HMM TIGR01549" /codon_start=1 /transl_table=11 /product="phosphonoacetaldehyde phosphonohydrolase" /protein_id="AAS40368.1" /db_xref="GI:42736433" 931. .2028 /locus_tag="BCE1440" 931. .2028 /locus_tag="BCE1440" /note="identified by match to protein family HMM PF00266" /codon_start=1 /transl_table=11 /product="aminotransferase, class v" /protein_id="AAS40369.1" /db_xref="GI:42736434" /translation="MTENHYLLLTTPGLTTTTRKVEVMLYDMCTWDVEYNTMVQDVRA KLVSLATKEEKYTTVLMQSGSTFVSVAIGSVIPKNGKLLVCTNGAYGKRVQWAEIM LHDVVVSTQTEEMEPNTNVVEVKLIQQDKEITHAIIVHCETTTGIINPIVDYKCLKGQ YKVLVDAMSSFGGIEDIDIAELQIDFLISSANKKCIQVPGFGFVIARKEDELLKCKGQ AKSLDLDYQDMQWENQKWRFTSPHVFHAFQALLELEKGGVGAARYNRYNNQ KLLVNRMEIGKPLVDYKQSPITTSFIYPEEVPDFPQLYNELKRDGFGVIYPYPGKIS VDTFRIGIDVHEADINRLVDSIARGVVG"
gene	2025. .3209 /locus_tag="BCE1441" 2025. .3209 /locus_tag="BCE1441" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein"	gene CDS	2025. .3209 /locus_tag="BCE1441" 2025. .3209 /locus_tag="BCE1441" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein"

Query Match	4.3%;	Score 50.6;	DB 1;	Length 20693;
Best Local Similarity	50.2%;	Pred. No. 0.0005;		
Matches 125;	Conservative 0;	Mismatches 124;	Indels 0;	Gaps 0;
382	CATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGGCGGATAACGGGCACGCCGG	441		
11667	CTTTTTCAGAAATTTCCGACGGCCCGCCGACAGCGGGTGTGACCCAGCGGTGCCGC	11608		
442	AGACTGAGCTTCAGCAGTACCATACCAACCGCTTCATTTCCGAGGCGATGACCACCA	501		
11607	ACAACAGCGATTCTACACAGCACATTGCCAACCTTCGTATCGAGCTGAGCACCAACA	11548		
502	CACGTGCAATCCGGTAGACGGTAACGCTGGGAAAAGGCGACCTGCCATTAACACATCTC	561		
11547	GGAGCGCATGGCGGATAAACGGATAGGATTGGCTTGGAAACCGAGGAACAGACCGCGT	11489		
562	CGCTCATATCCAGGTGTTCTGTCTGCTGACGACGACGTGCTTCGTATTCTTCACGCCCGG	621		
11487	CTGCGATACCAAGTCTGCCCGCAGACGCTTGACCTCGGCGACCGGGCGTCAATCCCGG	11428		
622	CGCCACCA	630		
11427	TGCCGATCA	11419		

RESULT 40	
AE017269/c	
LOCUS	294300 bp DNA linear BCT 23-FEB-2004
DEFINITION	Bacillus cereus ATCC 10987, section 6 of 18 of the complete genome.

```

/protein_id="AAS40370.1"
/db_xref="GI:42736435"
/translation="MKVFCGGGAKICREAILDLVQSFSEFITTAVDFNBEGLKVVYE
WLNDRPRVDFVQVNTNEDTVAKMGYDVIMDGTTIKINGLSTICIAEAGCHGVNLNG
PGENDSHSICFVONGKCLPCFGMTPGVTOAMMAANQOLHTVESVRVSHGSRPIAP
SASITETTYEDPRLPSRTYDEGEKQVPPFAPREIELPAPYKTKTOYIIPHSET
ITLAKALENGVKLIETHTGTPKQNMQLVRALYDYGILRNDQVEINGKEGIMDCIAK
YLLQSGQGETIYGIALHVEVIGMKNEKORRVLVYTHPLSDSGVVGWEKIRAYTRN
VGPFGIATELIAKGNVNVKGVVTPPEAFVKPQIIFDLKKRGIHIEEVFVEKESYN
FV"
gene
3250..4032
/locus_tag="BCE1442"
CDS
3250..4032
/locus_tag="BCE1442"
/notes="identified by match to protein family HMM PF00455"
/codon_start=1
/transl_table=11
/product="transcriptional regulator, Deor family"
/protein_id="AAS40371.1"
/db_xref="GI:42736436"
/translation="MSVVGEEKRTILEKVEFGKVKSVELAREFAVSTETIRYLEE
LDREKLUKVGAGVQVPGAGIEAPMLERMLHIEKKRIGYKAATFVEDGDVIAIDD
GSTPLQWVYLVRKNTITVTSFPVATOLISSINKMFHGEVLFPGKVPKHSRVS
GSIQOVIQOPHFKAQFVSDGLLPGRVSSFELEKAKLSEAMMKLAETKIYLCDHTK
VGKGNRIAPRSIQHVICDKMPSFEVEKHNQWTIC"
4119..4313
/locus_tag="BCE1443"
gene
4119..4313
/locus_tag="BCE1443"
/notes="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAS40372.1"
/db_xref="GI:42736437"
/translation="MDTFSHLPHIYQYKKSWSKERSMSERIYNKLVLYANLQKI
GVINEKKEIHLHTIGKAL"
4479..6746
/locus_tag="BCE1444"
gene
4479..6746
/locus_tag="BCE1444"
/notes="identified by match to protein family HMM PF00395;
match to protein family HMM PF00560"
/codon_start=1
/transl_table=11
/product="internalin, putative"
/protein_id="AAS40373.1"
/db_xref="GI:42736438"
/translation="MKALVAVATLIAIPPAAYSTPALAAIKIENQSVAAASERTYDE
IKIYKQKDEPSWSQIKDPKVTIADGKKIVVTMQSDSYFQYLRIEDKNQPGVPHD
VKVLSDEKNGTKVQFETGEFEKGNMQMHILIPAIGYDHKYQVQFEIKDPTVGDK
BTEKPDNNSNGNTETDHPVDNQNMITDNKLRLVKKVFNKDLNTPITKEELLQVK
DLPLNTNEILDYSALKYMPNLKSLTVANAKIKOPSFYTSKQLKHLALRGNEFSDVTP
IVKMDNLSDLSNNKLTNVAPLLEMKNVKSLYLSGNOIEDVTLAKWEOOLDYLNLAN
NKIYNTPLSALKNVTYLTLAGNIEDIKPLYSPLKDLVLRNKVKDLSGIEQMKQL
BELWIGNEIKNVTPLSKMQLKQLHLPNNELKIDITPLSSLVNQLKDLSEIYSDIT
PASNLKXLFVANSIRDVPRVIELSKYAYINVQKVFLEETEYNEKVPYIEK
DGKISTIRKLRESGTSDNDAKWSPTGEKYFEGVQDFADTGIFFTSVGIQNVVES
KEDNKSDEKVKVQEPKDPKGHSEBETHYLAKENIFKYGNGQRFQFGDRIPIRGQV
ASLVQRVYLKENYQKERTDITKGHMFEODIATVAOAGIMQGDGTGEFRPDKVLTZY
EMSVLYKVFOLKEDENKNVNDVPTGHWABEVKALADNNISKGDGKGNFLGDFFV
TREQYAQFLYNAINMK"
6996..7625
/locus_tag="BCE1445"
gene
6996..7625
/locus_tag="BCE1445"
/notes="identified by match to protein family HMM PF00596"
/codon_start=1
/transl_table=11
/product="class II aldolase/adducin domain protein"
/protein_id="AAS40374.1"
/db_xref="GI:42736439"
/translation="MLFLLKKNELKDVKSELALRDWPFYGTKISLSLCTSKEPLTFLV
```

```

NVRGRDKLSEEDFIVVNCMECPVFNENKPAAESFMHADIYKKSABECLIQVQTVD
SHLSELYGKEVITFDKRSYKVRVFGKEGIMTPIVEDEKPKFADLLENNVPNFSEG
GGVVLVHNYGMIVMGKTPPEYAKWLEGIEYLMNHYVLLMIKGAKSSVI"
7751..8374
/locus_tag="BCE1446"
gene
7751..8374
/locus_tag="BCE1446"
/notes="identified by match to protein family HMM PF00571"
/codon_start=1
/transl_table=11
/product="CBS domain protein"
/protein_id="AAS40375.1"
/db_xref="GI:42736440"
/translation="MRVKHPLPKQVTFCKMDSGKALQIMNETGFRAIPVLAEDE
KFMPIIYKDLLEKCKNGLELSEHMLEDSAAFPEKDSFFRAFYVIRLPLFLAV
LNDYNEFVGLTHSNITDVIEDSFGMRTGGYILTATQCKGTKEIGLTKLKAFFHIG
LFTLDGQDVIRRVINVTIDELNEKRLQLIEKIEKGFVRVSHVDYI"
complement(8442..9212)
/locus_tag="BCE1447"
gene
complement(8442..9212)
/locus_tag="BCE1447"
/notes="identified by match to protein family HMM PF04748"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAS40376.1"
/db_xref="GI:42736441"
/translation="MRKYTIALILITMFLPSFLPVBKAKHTNKVAIVIDDGNNMKG
TKMLSLPIPLTVAVMPFLPSTKEDAJAAHKHGEVITHMPMEPIKGKCEWLGPKAIT
TDLSDSEINNRLSOAQIEVPHAIENHNMGSKVTVADERIVRLILAAKCKHGLFYLDK
TNPKSVVPKIGELGVPILIEHQLFEDDVYTAHISKQALLIKLQKEKPINVAIGHVG
PPGEITSRVLETISIPKIREHADFIFLSDLALSPPPVSK"
9549..9749
/locus_tag="BCE1448"
gene
9549..9749
/locus_tag="BCE1448"
/notes="stage 0 sporulation regulatory protein,
degenerate"
/codon_start=1
/transl_table=11
/product="protein id="AAS40377.1"
/db_xref="GI:42736442"
/translation="MFTVKRYTLEKLSRDIHMKREMIQLGLTSLNSTETIQVSQE
LDKLIQYQCYKENKHQNGYQL"
complement(9859..11355)
/locus_tag="BCE1449"
gene
complement(9859..11355)
/locus_tag="BCE1449"
/EC_number="2.7.3.-"
/notes="identified by match to protein family HMM PF00512;
match to protein family HMM PF02518"
/codon_start=1
```

```

Query Match      4.38; Score 50.6; DB 1; Length 294300;
Best Local Similarity 50.28; Pred. No. 0.00089;
Matches 125; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy - 255 AACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATATCCGGAGCGTTCGGTCGG 314
Db 183745 AACTGTCGTAACACATCTGCTTGTCTCGCTCTCCCATATACGCTGAAGTTCTTCATCC 183686

Qy 315 GATAAAATTCGGCAGTCGGCGGTCCTCATGAGACATCCCCCAGCGGTAAACAGCGTC 374
Db 183685 TTTAATAGCTGAATAGCTTGTATTGGCCACTCTCTGTATGCGCAACTTCACATAAATAT 183626

Qy 375 CCTGTACATCTTCTCAATGACATCAGGATCCCGCCGCTCTCACTGGCGGATACCGGC 434
Db 183625 CCTGTATCAGATGTTGAATGACTTCTGGAATACCTCAACCCCTGTGTTCCGATCAAGGT 183566

Qy 435 ACGCCGAGACTGACCGTTCAGCGGTACCATACCAACCGCTTCATTTTCCGAAGGCATG 494
Db 183565 ACACCACGCGCATCGCTTCTTAATAAACAAGACCAAAACTCTCTTCTCTGATAGAAGT 183506
```

Qy 495 ACCACCACA 503
| | | | |
Db 183505 AACATTAAA 183497

Search completed: June 4, 2005, 14:02:27
Job time : 5415 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 07:29:15 ; Search time 728 Seconds
(without alignments)
9603.317 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 cfcagagatgagaaaaa.....ttttactttttctctgcag 1181

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseqn_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1181	100.0	1181	3	Aaz36102 Nucleic a
C 2	159.4	13.5	1642	2	Aaz36102 Nucleic a
C 3	64.2	5.4	1134	13	Adt48421 Bacterial
4	64.2	5.4	10181	12	Adn36367 Plasmid p
5	64.2	5.4	10801	4	Aas02323 Plasmid p
C 6	57.2	4.8	1197	8	Aca38556 Prokaryot
C 7	57.2	4.8	1200	8	Aca40610 Prokaryot
8	57.2	4.8	110000	4	Continuation (25 o
9	57.2	4.8	110000	4	Continuation (25 o
10	55	4.7	110000	12	Continuation (12 o
C 11	55	4.7	110000	12	Continuation (10 o
12	55	4.7	110000	12	Continuation (12 o
C 13	55	4.7	110000	12	Continuation (10 o
14	55	4.7	110000	12	Continuation (12 o
C 15	55	4.7	110000	12	Continuation (10 o
16	50.4	4.3	1202	13	Adn47960_09
C 17	50.2	4.3	1143	8	Adn47960_09
C 18	49	4.1	1135	6	Adn47960_09
C 19	48.6	4.1	783	6	Abk78827 Bacillus
C 20	48.6	4.1	1179	8	Aca27126 Prokaryot

C 21	48.2	4.1	1113	13	Adt45647 Bacterial
C 22	48.2	4.1	1146	13	Adn46591 Bacterial
23	46.6	3.9	569	9	Adn46591 Bacterial
24	46.6	3.9	609	9	Adn46591 Bacterial
25	46.6	3.9	615	9	Adn46591 Bacterial
26	46.6	3.9	619	9	Adn46591 Bacterial
27	46.6	3.9	640	9	Adn46591 Bacterial
28	46.6	3.9	647	9	Adn46591 Bacterial
29	46.6	3.9	651	9	Adn46591 Bacterial
30	46.6	3.9	669	9	Adn46591 Bacterial
31	46.6	3.9	683	9	Adn46591 Bacterial
32	46.6	3.9	693	9	Adn46591 Bacterial
33	46.6	3.9	713	9	Adn46591 Bacterial
34	46.2	3.9	570	11	Abd06478 Pseudomon
C 35	46.2	3.9	714	11	Abd06478 Pseudomon
C 36	46.2	3.9	783	6	Abk78820 Bacillus
C 37	46.2	3.9	1251	11	Abd06412 Pseudomon
C 38	46.2	3.9	1584	11	Abd06505 Pseudomon
C 39	45.6	3.9	1119	10	Abd06505 Pseudomon
C 40	45.6	3.9	69727	10	Abd06505 Pseudomon
C 41	45.6	3.9	110000	10	Abd06505 Pseudomon
C 42	44.4	3.8	566	9	Adn46591 Bacterial
C 43	44.2	3.7	1120	13	Adn46591 Bacterial
C 44	42.8	3.6	1155	8	Adn46591 Bacterial
C 45	41.6	3.5	1227	6	Abk74480 Bacillus

ALIGNMENTS

RESULT 1

Aaz36102

ID AAZ36102 standard; DNA; 1181 BP.

XX

AC AAZ36102;

DT 11-FEB-2000 (first entry)

XX

DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX

KW Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

KW enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.

XX

OS Escherichia coli.

XX

PN WO955908-A2.

XX

PD 04-NOV-1999.

XX

PF 27-APR-1999; 99WO-FR001000.

XX

PR 28-APR-1998; 98FR-00005329.

XX

PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PI

PI Frechon DTM, Laure FC, Thierry D;

DR

DR WPI; 2000-013443/01.

PT

PT New nucleic acid containing sequences specific to enterohemorrhagic

PT

PT Escherichia coli, particularly serotype O157:H7, used for detecting these

PT

PT bacteria in food.

XX

XX Claim 1; Fig 2; 48pp; French.

PS

PS The present sequence is specific to enterohemorrhagic Escherichia coli

CC

CC (EHEC). The sequence associated with the presence of virulence factors

CC

CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have

CC

CC 68% homology with the virK gene which codes for virulence proteins of

CC

CC Shigella flexneri. The present sequence is of plasmid origin. Fragments

CC

CC of the present sequence are used, as probes and primers, for detection of

CC

CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or

CC

CC animal samples, foods or the environment. The fragments are also useful

PT Producing B6 vitamer such as pyridoxine, pyridoxal or pyridoxamine,
PT involves culturing organism with increased Yaad and/or YaaE activity as
PT compared to parent organism.

XX Example 4; SEQ ID NO 8; 75pp; English.

XX The invention relates to a method of producing (M1) a B6 vitamer, which
CC involves culturing an organism (I) with increased Yaad and/or YaaE
CC activity, or Epd, PdxA, PdxJ, PdxP, PdxX, PdxH and/or Dxs activity as
CC compared to the parent organism, or a microorganism that overexpresses
CC Bacillus B6 vitamer biosynthetic gene, under conditions such that B6
CC vitamer is produced. (M1) is useful for producing B6 vitamer such as
CC pyridoxine, pyridoxal and pyridoxamine. This sequence represents the
CC plasmid pDX17R which contains the Bacillus subtilis yaadE operon used to
CC express these proteins.

XX Sequence 10181 BP; 2655 A; 2363 C; 2041 G; 3122 T; 0 U; 0 Other;

Query Match 5.4%; Score 64.2; DB 12; Length 10181;
Best Local Similarity 54.4%; Pred. No. 3.1e-09;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 335 GCGGTCCATGCGAGACATCCCGCAGGTAAACAGCGTCCGTGTACATCTTCTGCAAT 394
DB 5870 GGCTGTGGCGCGGTAAACATCCCGACATCCACCAAAATCCGCTCACATTTGTTTTAT 5929
QY 395 GACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGACGCGGAGACTGACGCTTC 454
DB 5930 AACCTAGGATACCGCAATGTTGTTCCAAATACAGGACATCCCGACGCTTCGCTTC 5989
QY 455 AGCCAGTACCATACCAACGCTTCATTTTCCGAAGCATGACCAACCACTGGCAATCCG 514
DB 5990 AAGCAGGACAAAGCCAAAGCTTTCTTTTTCAGATAGCAGCAGCTTCAAATCGCTAATAGA 6049
QY 515 GTAGACGGTAACGCTGGGAAAGGACCTGCGATTAAACATCTCCGCTCATTC 571
DB 6050 ATAAAGATCTTCAACACGGCTTTGATTTTCAAGCATTTAAGACTTGGTCTTTCCAAGCC 6106

RESULT 5

AAS02323
ID AAS02323 standard; DNA; 10801 BP.

AC AAS02323;

XX 06-AUG-2003 (revised)

DT 18-JUL-2001 (first entry)

XX Plasmid pAN240 carrying the B. subtilis PanBCD operon.

XX PanBCD operon; pantothenate biosynthesis; cyclic; circular; pAN240;
KW vitamin B5; nutritional supplement; panto-compound; pantoate; ds.

XX Bacteriophage.

OS Escherichia coli.

OS Bacillus subtilis.

OS Chimeric.

XX WO200121772-A2.

PN 29-MAR-2001.

XX 21-SEP-2000; 2000WO-US025993.

XX 21-SEP-1999; 99US-00400494.

PR 07-JUN-2000; 2000US-0210072P.

PR 28-JUL-2000; 2000US-0221836P.

PR 24-AUG-2000; 2000US-0227860P.

XX (OMNI-) OMNIGENE BIOPRODUCTS.

XX Yocum RR, Patterson TA, Hermann T, Pero JG;

XX
XX

DR WPI; 2001-218644/22.

XX New recombinant microorganism which overexpress a Bacillus subtilis
PT pantothenate biosynthetic enzyme, useful for the high yield production of
PT panto-compounds such as pantothenate and pantoate.

XX Example 1; Page 232-238; 292pp; English.

XX The sequence, PAN240, is a plasmid carrying the B. subtilis PanBCD operon
CC which encodes three enzymes of the pantothenate biosynthetic pathway
CC (PanB, C and D). Pantothenate, also known as vitamin B5, is used as a
CC nutritional supplement in mammals and humans. The invention concerns
CC methods of producing recombinant microorganisms overexpressing at least
CC one B. subtilis pantothenate biosynthetic enzyme. The microorganisms and
CC methods of producing them are useful for producing a panto-compound such
CC as pantothenate or pantoate, which is a nutritional requirement for
CC livestock and humans. The methods are also useful for the identification
CC of pantothenate kinase modulators. Panto-compounds are produced at a
CC significantly higher yield than prior art methods and can be produced
CC independent of the need to feed precursors which decreases expense.
CC (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 10801 BP; 2797 A; 2486 C; 2151 G; 3367 T; 0 U; 0 Other;

Query Match 5.4%; Score 64.2; DB 4; Length 10801;
Best Local Similarity 54.4%; Pred. No. 3.2e-09;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 335 GCGGTCCATGCGAGACATCCCGCAGGTAAACAGCGTCCGTGTACATCTTCTGCAAT 394
DB 6490 GGCTGTGGCGCGGTAAACATCCCGACATCCACCAAAATCCGCTCACATTTGTTTTAT 6549
QY 395 GACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGACGCGGAGACTGACGCTTC 454
DB 6550 AACCTAGGATACCGCAATGTTGTTCCAAATACAGGACATCCCGACGCTTCGCTTC 6609
QY 455 AGCCAGTACCATACCAACGCTTCATTTTCCGAAGCATGACCAACCACTGGCAATCCG 514
DB 6610 AAGCAGGACAAAGCCAAAGCTTTCTTTTTCAGATAGCAGCAGCTTCAAATCGCTAATAGA 6669
QY 515 GTAGACGGTAACGCTGGGAAAGGACCTGCGATTAAACATCTCCGCTCATTC 571
DB 6670 ATAAAGATCTTCAACACGGCTTTGATTTTCAAGCATTTAAGACTTGGTCTTTCCAAGCC 6726

RESULT 6

ACA38556/c

ID ACA38556 standard; DNA; 1197 BP.

XX ACA38556;

XX 19-JUN-2003 (first entry)

XX Prokaryotic essential gene #20213.

DE Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.

XX Mycobacterium bovis.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (BLIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX
XX WPI: 2003-029926/02.
DR P-PSDB; ABU34686.
XX
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX
XX Claim 14; SEQ ID NO 26426; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid;
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 1197 BP; 171 A; 388 C; 429 G; 209 T; 0 U; 0 Other;

Query Match 4.8%; Score 57.2; DB 8; Length 1197;
Best Local Similarity 54.2%; Pred. No. 1.8e-07;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 272 TTCTCTGGCGCTGGCTCCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGGCGAG 331
DB 1119 TTCTCGCGCGGGCGGCCATCGCGAGCGGCCGATCCGATCATGACGACATCGGCCAC 1060

QY 332 TGGCGCGGTCCATGACAGACATATCCCGGTAACAGCGTCCCTGTACATCTTCGTG 391
DB 1059 GGGCTCGGCCACCGCGTCCACCGACCTACCGTCCGACCACTAGCCAGTCTTGTGTGCTG 1000

QY 392 AATGACATCAGGATCCCGCGCTCTCATCTGGGATACCGGACCGCGAGACTGACGC 451
DB 999 CACCGTTTCGGCGGCTCCGCGAGAAATGCGCGGATATACCGGACCGCGCGCGGAGGC 940

QY 452 TTGACGCGATACCATACCAACGCTTTCATTTTCC 485
DB 939 TTCAGGAACACGATGCGCAGCCCTGACGTCC 906

RESULT 7

ACA40610/c

ID ACA40610 standard; DNA; 1200 BP.

XX

AC ACA40610;

XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #22267.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX
XX Mycobacterium tuberculosis.
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX 06-SEP-2001; 2001US-00948993.
XX 25-OCT-2001; 2001US-0342923P.
XX 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI: 2003-029926/02.
DR P-PSDB; ABU36740.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 28480; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid;
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1200 BP; 172 A; 388 C; 430 G; 210 T; 0 U; 0 Other;

Query Match 4.8%; Score 57.2; DB 8; Length 1200;
Best Local Similarity 54.2%; Pred. No. 1.8e-07;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Db	47373	TTCTCGCGCGCGCCATCGCGACGGCCCGATCCGATCGATCAGCAACTCGGCCAC	47432
Qy	332	TGCGCCGCTCATGACAGACATCCCCACGGGTAACAGCGTCCCTGTCACTTCTGTG	391
Db	47433	GGCGTCGGCCACCCCGCTCCACCGACCTACCGTCGACCACTAGCCCACTTGTGTGCTG	47492
Qy	392	AATGACATCAGGATCCGCGCCGCTCACTCGGCGATACGGGACGCGGAGACTGACGC	451
Db	47493	CACCGTTTCGGCGCTCCGCCAGAAATGCGCGCGATTACCGGCACGCGCGCGGAGGC	47552
Qy	452	TTACGCCAGTACCATACCAACGCTTCAATTTCC	485
Db	47553	TTGAGGAGACAGATGCCCAAGCCTCGACGTCC	47586
RESULT 10			
ADN46845_11			
Continuation (12 of 21) of ADN46845 from base 1100001 (Thermococcus kodakaraensis KOD1 9			
WP Sequence split into 21 fragments LOCUS ADN46845 Accession Adn46845			
WP	Fragment Name	Begin	End
WP	ADN46845_00	1	110000
WP	ADN46845_01	100001	210000
WP	ADN46845_02	200001	310000
WP	ADN46845_03	300001	410000
WP	ADN46845_04	400001	510000
WP	ADN46845_05	500001	610000
WP	ADN46845_06	600001	710000
WP	ADN46845_07	700001	810000
WP	ADN46845_08	800001	910000
WP	ADN46845_09	900001	1010000
WP	ADN46845_10	1000001	1110000
WP	ADN46845_11	1100001	1210000
WP	ADN46845_12	1200001	1310000
WP	ADN46845_13	1300001	1410000
WP	ADN46845_14	1400001	1510000
WP	ADN46845_15	1500001	1610000
WP	ADN46845_16	1600001	1710000
WP	ADN46845_17	1700001	1810000
WP	ADN46845_18	1800001	1910000
WP	ADN46845_19	1900001	2010000
WP	ADN46845_20	2000001	2089378
Query Match 4.7%; Score 55; DB 12; Length 110000;			
Best Local Similarity 53.5%; Pred. No. 1.4e-05;			
Matches 115; Conservative 0; Mismatches 100; Indels 0; Gaps 0;			
Qy	264	TCAATATCTTCTCTGGCGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAA	323
Db	21802	TCAACGGCTTTCTCCCGCTTCTCCAAACTTTTGGCAAGCCCTTTGTGCGAGAGGAGC	21743
Qy	324	TGCGCGAGTGGCGGTCATGCGAGACACATCCCCCAGGGTAACAGCGTCCCTGTGACA	383
Db	21742	TTTAAGACTGCTTCGGCAAGTGCCTTCTGTCGCCAGGGGGAGAGTATGCCGCTTTCA	21683
Qy	384	TTCTTCTGAATGATCAGGATCCCGCGCTCTCACTGCGATACGGGCACGCCGGAG	443
Db	21682	CTTCTTTAAGACTTCAGGATCCCAACACCGTTGTCTGTGACAACTGGAATTCAGAG	21623
Qy	444	ACTGACGTTTCAGCCAGTACCATACCAACGCTTC	478
Db	21622	GCCATCGCTCAAGCACTACTATCCCAAGGCCTC	21588
RESULT 12			
ADN46123_11			
Continuation (12 of 21) of ADN46123 from base 1100001 (Thermococcus kodakaraensis KOD1			
WP Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123			
WP	Fragment Name	Begin	End
WP	ADN46123_00	1	110000
WP	ADN46123_01	100001	210000
WP	ADN46123_02	200001	310000
WP	ADN46123_03	300001	410000
WP	ADN46123_04	400001	510000
WP	ADN46123_05	500001	610000
WP	ADN46123_06	600001	710000
WP	ADN46123_07	700001	810000
WP	ADN46123_08	800001	910000
WP	ADN46123_09	900001	1010000
WP	ADN46123_10	1000001	1110000
WP	ADN46123_11	1100001	1210000
WP	ADN46123_12	1200001	1310000
WP	ADN46123_13	1300001	1410000
WP	ADN46123_14	1400001	1510000
WP	ADN46123_15	1500001	1610000
WP	ADN46123_16	1600001	1710000
WP	ADN46123_17	1700001	1810000
WP	ADN46123_18	1800001	1910000
WP	ADN46123_19	1900001	2010000
WP	ADN46123_20	2000001	2089378
Query Match 4.7%; Score 55; DB 12; Length 110000;			
Best Local Similarity 53.5%; Pred. No. 1.4e-05;			
Matches 115; Conservative 0; Mismatches 100; Indels 0; Gaps 0;			
Qy	264	TCAATATCTTCTCTGGCGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAA	323
Db	67576	TCAACGGCTTTCTCCCGCTTCTCCAAACTTTTGGCAAGCCCTTTGTGCGAGAGGAGC	67635
Qy	324	TGCGCGAGTGGCGGTCATGCGAGACACATCCCCCAGGGTAACAGCGTCCCTGTGACA	383
Db	67636	TTTAAGACTGCTTCGGCAAGTGCCTTCTGTCGCCAGGGGGAGAGTATGCCGCTTTCA	67695
Qy	384	TTCTTCTGAATGATCAGGATCCCGCGCTCTCACTGCGATACGGGCACGCCGGAG	443
Db	67696	CTTCTTTAAGACTTCAGGATCCCAACACCGTTGTCTGTGACAACTGGAATTCAGAG	67755
Qy	444	ACTGACGCTTCAGCCAGTACCATACCAACGCTTC	478
Db	67756	GCCATCGCTCAAGCACTACTATCCCAAGGCCTC	67790
RESULT 11			
ADN47591_09/c			
Continuation (10 of 21) of ADN47591 from base 900001 (Thermococcus kodakaraensis KOD1 9			
WP Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591			
WP	Fragment Name	Begin	End
WP	ADN47591_00	1	110000
WP	ADN47591_01	100001	210000
WP	ADN47591_02	200001	310000
WP	ADN47591_03	300001	410000
WP	ADN47591_04	400001	510000
Query Match 4.7%; Score 55; DB 12; Length 110000;			
Best Local Similarity 53.5%; Pred. No. 1.4e-05;			
Matches 115; Conservative 0; Mismatches 100; Indels 0; Gaps 0;			
Qy	264	TCAATATCTTCTCTGGCGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAA	323
Db	67576	TCAACGGCTTTCTCCCGCTTCTCCAAACTTTTGGCAAGCCCTTTGTGCGAGAGGAGC	67635

QY	324	TCGCGCAGTCCGCGCTCCATGAGACACATCCCCACGGGTAAACGCGTCCCTGTCA	383
Db	67636	TTTAAAGACTGCTTCGGCAAGTCCGCTTCGTGCGCAGGGGACGAGTATGCGCGTTTCA	67695
QY	384	TTCTTCTGATGACATCAGGATCCCGCGTCTCACTGCGGATACGGGCACGCCGGAG	443
Db	67696	CTTCTTTAACGACTTCAGGGATCCCAACCGTTGTGCTGCAACTGGAATTCAGAG	67755
QY	444	ACTGAGCGCTTCAGCCAGTACCATACCAACGCTTC	478
Db	67756	GCCATCGCCTCAGCACTACTATCCCAAGGCCTC	67790
RESULT 13			
ADN47209_09/c			
Continuation (10 of 21) of ADN47209 from base 900001 (Thermococcus kodakaraensis KOD1 ge			
WP	Sequence split into 21 fragments		
WP	Fragment Name	Begin	End
WP	ADN47209_00	1	110000
WP	ADN47209_01	100001	210000
WP	ADN47209_02	200001	310000
WP	ADN47209_03	300001	410000
WP	ADN47209_04	400001	510000
WP	ADN47209_05	500001	610000
WP	ADN47209_06	600001	710000
WP	ADN47209_07	700001	810000
WP	ADN47209_08	800001	910000
WP	ADN47209_09	900001	1010000
WP	ADN47209_10	1000001	1110000
WP	ADN47209_11	1100001	1210000
WP	ADN47209_12	1200001	1310000
WP	ADN47209_13	1300001	1410000
WP	ADN47209_14	1400001	1510000
WP	ADN47209_15	1500001	1610000
WP	ADN47209_16	1600001	1710000
WP	ADN47209_17	1700001	1810000
WP	ADN47209_18	1800001	1910000
WP	ADN47209_19	1900001	2010000
WP	ADN47209_20	2000001	2089378
Query Match			
Best Local Similarity 4.7%; Score 55; DB 12; Length 110000;			
Matches 115; Conservative 0; Mismatches 100; Indels 0; Gaps 0;			
QY	264	TCAATATCTTCTCTGGCGCTGCGCATCATCCGGAAGCGTTCCGGTCGGGATAAAAA	323
Db	67576	TCAACGGCTTTTCTCCCGGCTTCTCCAAACTTTTGGCAAGCCCTTTGTGCGAGAGGAGC	67635
QY	324	TCGCGCAGTCCGCGCTCCATGAGACACATCCCCACGGGTAAACGCGTCCCTGTCA	383
Db	67636	TTTAAAGACTGCTTCGGCAAGTCCGCTTCGTGCGCAGGGGGACGAGTATGCCGCTTTC	67695
QY	384	TTCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGGCACGCCGGAG	443
Db	67696	CTTCTTTAACGACTTCAGGGATCCCAACCGTTGTGCTGCAACTGGAATTCAGAG	67755
QY	444	ACTGAGCGCTTCAGCCAGTACCATACCAACGCTTC	478
Db	67756	GCCATCGCCTCAGCACTACTATCCCAAGGCCTC	67790
RESULT 15			
ADN47960_09/c			
Continuation (10 of 21) of ADN47960 from base 900001 (Thermococcus kodakaraensis KOD1 g			
WP	Sequence split into 21 fragments		
WP	Fragment Name	Begin	End
WP	ADN47960_00	1	110000
WP	ADN47960_01	100001	210000
WP	ADN47960_02	200001	310000
WP	ADN47960_03	300001	410000
WP	ADN47960_04	400001	510000
WP	ADN47960_05	500001	610000
WP	ADN47960_06	600001	710000
WP	ADN47960_07	700001	810000
WP	ADN47960_08	800001	910000
WP	ADN47960_09	900001	1010000
WP	ADN47960_10	1000001	1110000
WP	ADN47960_11	1100001	1210000
WP	ADN47960_12	1200001	1310000
WP	ADN47960_13	1300001	1410000
WP	ADN47960_14	1400001	1510000
WP	ADN47960_15	1500001	1610000
WP	ADN47960_16	1600001	1710000
WP	ADN47960_17	1700001	1810000
WP	ADN47960_18	1800001	1910000
WP	ADN47960_19	1900001	2010000
WP	ADN47960_20	2000001	2089378
Query Match			
Best Local Similarity 4.7%; Score 55; DB 12; Length 110000;			
Matches 115; Conservative 0; Mismatches 100; Indels 0; Gaps 0;			
QY	264	TCAATATCTTCTCTGGCGCTGCGCATCATCCGGAAGCGTTCCGGTCGGGATAAAAA	323
Db	21802	TCAACGGCTTTTCTCCCGGCTTCTCCAAACTTTTGGCAAGCCCTTTGTGCGAGAGGAGC	21743
QY	324	TCGCGCAGTCCGCGCTCCATGAGACACATCCCCACGGGTAAACGCGTCCCTGTCA	383
Db	21742	TTTAAAGACTGCTTCGGCAAGTCCGCTTCGTGCGCAGGGGGACGAGTATGCCGCTTTC	21683
QY	384	TTCTTCTGATGACATCAGGATCCCGCGCTCTCACTGCGGATACGGGCACGCCGGAG	443
Db	21682	CTTCTTTAACGACTTCAGGGATCCCAACCGTTGTGCTGCAACTGGAATTCAGAG	21623
QY	444	ACTGAGCGCTTCAGCCAGTACCATACCAACGCTTC	478
Db	21622	GCCATCGCCTCAGCACTACTATCCCAAGGCCTC	21588
RESULT 14			
ADN46464_11			
Continuation (12 of 21) of ADN46464 from base 1100001 (Thermococcus kodakaraensis KOD1 g			
WP	Sequence split into 21 fragments		
WP	Fragment Name	Begin	End
WP	ADN46464_00	1	110000
WP	ADN46464_01	100001	210000
WP	ADN46464_02	200001	310000
WP	ADN46464_03	300001	410000
WP	ADN46464_04	400001	510000
WP	ADN46464_05	500001	610000
WP	ADN46464_06	600001	710000
Query Match			
Best Local Similarity 4.7%; Score 55; DB 12; Length 110000;			
Matches 115; Conservative 0; Mismatches 100; Indels 0; Gaps 0;			
QY	264	TCAATATCTTCTCTGGCGCTGCGCATCATCCGGAAGCGTTCCGGTCGGGATAAAAA	323
Db	21802	TCAACGGCTTTTCTCCCGGCTTCTCCAAACTTTTGGCAAGCCCTTTGTGCGAGAGGAGC	21743
QY	324	TCGCGCAGTCCGCGCTCCATGAGACACATCCCCACGGGTAAACGCGTCCCTGTCA	383

Db 21742 TTTAAGACTCCCTCGGAGTCCGCTTGTGTCGCCAGGGGAGGAGTATGCCGCTTTCA 21683
 QY 384 TTCTTCTGAATGATCATCAGGAGTCCCGCCGCTCTCACTGCGGATACGGGCGAGCGCGGAG 443
 Db 21682 CTTTCTTTAAGACTTCAGGGATCCCAACCAACGTTGTGCTGACAACTGGAATTCAGAG 21623
 QY 444 ACTGAGCGCTTCAGCCAGTACCATACCAAGCGCTTC 478
 Db 21622 GCCATCGCTCAAGCACTACTATCCCAAGGCCTC 21588

RESULT 16
 ADS56895/c
 ID ADS56895 standard; cDNA; 1202 BP.
 AC ADS56895;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polynucleotide #8882.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.
 XX
 OS Bacteria.

XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 20-FEB-2003; 2003US-00369493.
 XX
 PR 21-FEB-2002; 2002US-0360039P.
 XX
 PA (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX
 DR WPI; 2004-061375/06.

XX
 PT New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 PS Claim 1; SEQ ID NO 32569; 122pp; English.

XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX
 SQ Sequence 1202 BP; 224 A; 378 C; 336 G; 264 T; 0 U; 0 Other;
 Query Match 4.3%; Score 50.4; DB 13; Length 1202;
 Best Local Similarity 51.3%; Pred. No. 3.3e-05;
 Matches 117; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 274 CTCTGGCGTGTGCTGCCATCATCCGGAAGCGTTCGGTGGGATATAAATCGGCAGTG 333
 Db 1104 CTGCGCCACGGGCGGCTCGCAGGCGAGACTCGGCGAGAGAACGCGAGGCTGTTGAGCAGGG 1045
 QY 334 CGCGCGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCTGTTCACATTCTTCTGAA 393
 Db 1044 CATGAGCGAGTGTCTCGGACATCACCTGCGCGGACCAAGGCGGTTGAGTCTGCTGGA 985
 QY 394 TGACATCAGGATCCCGCCGCTCTCACTGGCGGATACGGGCGAGACTGACGCTT 453
 Db 984 CAACTTCGCGAGGCCACCAAGTATTCGAGACGATCACTGGCGAGCGCGGCAATGCTT 925
 QY 454 CAGCGAGTACCATACCAAGCGTTCATTTTCGAAGGCGATGACCACA 501
 Db 924 CCAGGGCCCAATGCCAAATGGTTCGTACAGCGAAGGAGAGCGGCAA 877

RESULT 17
 ACA22859/c
 ID ACA22859 standard; DNA; 1143 BP.
 XX
 AC ACA22859;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Prokaryotic essential gene #4516.
 XX
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX
 OS Bacillus anthracis.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.

XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haeselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR P-PSDB; ABU18989.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 14; SEQ ID NO 10729; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression


```
ABK78827/c
ID ABK78827 standard; DNA; 783 BP.
XX
AC ABK78827;
XX
AT 13-AUG-2002 (first entry)
XX
DE Bacillus clausii genomic sequence tag (GST) #1670.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus clausii.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
PR 06-OCT-2000; 2000US-00680598.
PR 27-MAR-2001; 2001US-0279526P.
XX
PA (NOVO ) NOVOZYMES BIOTECH INC.
PA (NOVO ) NOVOZYMES AS.
XX
XX Berka R, Clausen IG;
XX
XX WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
PS Claim 11; SEQ ID NO 6118; 200pp; English.
XX
XX The invention describes a method of monitoring differential expression of
XX genes in a first Bacillus cell relative to expression of the genes in
XX other Bacillus cells, comprising hybridising labelled nucleic acid probes
XX isolated from Bacillus cells to a substrate containing array of Bacillus
XX genomic sequenced tags (GST), examining the array, and determining
XX relative gene expression by an observed hybridisation reporter signal of
XX a spot in the array. The method is useful for measuring the expression of
XX genes in a first Bacillus cell relative to expression of the same genes
XX in one or more second Bacillus cells. The method is useful for monitoring
XX global expression of several genes from a Bacillus cell, discovering new
XX genes, identifying possible functions of unknown open reading frames and
XX monitoring gene copy number variation and stability. Monitoring changes
XX in expression of genes may be used to provide a representation of the way
XX in which Bacillus cells adapt to changes in culture conditions,
XX environmental stress or other physiological provocation. Extensive follow
XX -up characterisation is unnecessary when one spot on an array equals one
XX gene or one open reading frame, since sequence information is available.
XX This sequence represents a genomic sequence tag (GST) used in the method
XX of the invention. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 783 BP; 225 A; 149 C; 199 G; 208 T; 0 U; 2 Other;

Query Match 4.1%; Score 48.6; DB 6; Length 783;
Best Local Similarity 49.8%; Pred. No. 0.0001;
Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy 354 TCCCCACGGGTACAGGTCCTTCATCTTCTGATGATGACATCAGGATCCGCC 413
Dy |||||
Dy 631 TTCCCAAGGGGGCATAATATACCAAGTTACTTCGTCTTCAACAACCTCCGGATGCCACCA 572
Qy 414 GTCTCACTGGCGATAACGGGACCGGAGACTGACGCTTCAGCGCAGTACCATACCAAC 473
Dy |||||
Dy 571 ATTTTGTGGGTACAGGTTACCCCGCATCGCTTCAGCGCAGTAAACCAAG 512

Oy 474 GCTTCATTTTCCGAGGCGATGACACCACTGGCAATCGGTAGACCGGTAGCGTGGG 533
Dy |||||
Dy 511 CTTTCTTTTTCGCTTAACCAACAGCATTTAAATCGCTCATAGACAACAAATTCAGCAATGC 452
Qy 534 AAAAGGGCACCTGCCATTAAACACATCTCGCTCATTTCCAGAGTGTTCTGTCTGTACGC 593
Dy |||||
Dy 451 TTTTGGTTCCTAGCATATAGGACATCGTTTGTCAGCGCGTTTCTTCCACAACATGCTC 392
Qy 594 AGACGTG 600
Dy |||
Dy 391 GCAATTG 385

RESULT 20
ACA27126/c
ID ACA27126 standard; DNA; 1179 BP.
XX
XX ACA27126;
XX
AT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #8783.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
XX Bordetella pertussis.
XX
XX WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlson KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX P-PSDB; ABU23256.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 14996; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
```

CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 1179 BP; 178 A; 395 C; 406 G; 200 T; 0 U; 0 Other;

Query Match 4.1%; Score 48.6; DB 8; Length 1179;
Best Local Similarity 55.7%; Pred. No. 0.00013;
Matches 93; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 326 GCGCAGTGGCCGGTCCATCGACACATCCGCCACGGGTAACAGCGTCCCTGTGCACATT 385
DB 999 GCGCAGGCCCTCGAGCGCAGCGCGGTCTGTGGAGGAGACAGAGCCCCGTCTGGCC 940
QY 386 CTTCTGAATGACATCAGGGATCCGCCCGTCTCACTGGCGATACCGGCACGCCGAGAC 445
DB 939 GTTGGCGAACATCTCGTGTAGCGCGCGGACATCCACGCCGACCCGCGAAGCGCACGC 880
QY 446 TGACGCTTCAGCCAGTACCATACCAACGCTTTCATTTTCGAGGGCA 492
DB 879 CGCGGCGCTCGACAAAGACCGTGCCTCCGACGCGCTCTTGGCGGTGGCA 833

RESULT 21
ADT45647/c
ID ADT45647 standard; cDNA; 1113 BP.

XX
AC ADT45647;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #20398.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

PT for expression of a polynucleotide encoding a polypeptide from a

PT microbial source, useful for producing plants with improved properties.
XX Claim 1; SEQ ID NO 44085; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 1113 BP; 379 A; 204 C; 242 G; 288 T; 0 U; 0 Other;

Query Match 4.1%; Score 48.2; DB 13; Length 1113;
Best Local Similarity 59.0%; Pred. No. 0.00017;
Matches 102; Conservative 0; Mismatches 68; Indels 3; Gaps 1;
QY 352 CATCCCCCACGGGTAAACGGTCCCTGTGCACATTCTTCTGAATGACATCAGGATCCCGC 411
DB 976 CATTTCCGGGTGGAACATAAAGCCCTGCACGTCTTTCTTTTATAACCTCAGGTATACCAC 917
QY 412 CCGTCTCACTGGCGATACGGGCACGCCGAGACTGACGCTTCAGCCAGTACCATACCAA 471
DB 916 CAACGCTCTGTAGCTATATATGGGAATCCCGAGGCCATGGCCCTCAAGGATTTACTATGCCGA 857
QY 472 AGCTTTTCATTTTC---CGAAGCATGACACCAACACACTGGCAATCCGGTAGACC 521
DB 856 AGCTTCCGATGATATCGAAGGTAGACGACGACGTCGCGCCATTCTGAAACC 804

RESULT 22

ADSA46591/c

ID ADSA46591 standard; cDNA; 1146 BP.

XX ADSA46591;

XX 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #1334.

XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
DR
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 25021; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1146 BP; 388 A; 207 C; 253 G; 298 T; 0 U; 0 Other;
Query Match 4.1%; Score 48.2; DB 13; Length 1146;
Best Local Similarity 59.0%; Pred. No. 0.00017;
Matches 102; Conservative 0; Mismatches 58; Indels 3; Gaps 1;
Qy 352 CATCCGCCGGAACAGCGTCCCTGTCTCATATTCTTCTGAATGACATCAGGATCCCGC 411
Db 985 CATTCGGGGTGGAACTAAAAGCCCTGCACGTGTTTCTTTATTAACCTCAGGTATACCA 926
Qy 412 CCGTCTCACTGGGATACGGGACGCGGAGACTGACGCTTACGCCAGTACCATACCAA 471
Db 925 CAACGCTGTAGCTATTAATGGAACTCCGAGGCGCATGGGCTCAAGGATTACTATGCCGA 866
Qy 472 ACGTTTCAATTTC---CGAAGGAGATGACACCACTGCGCAATCCGGTAGACC 521
Db 865 ACGTTCCGATGATATCGAAGGTAAAGACGACAGCTGGGCATCTTGAAACC 813
RESULT 23
ACLI17707
ID ACLI17707 standard; DNA; 569 BP.
XX
AC ACLI17707;
XX
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7698.

XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Akashinriki).
XX WO2003057877-A1.
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
PR 20-DEC-2001; 2001JP-00387131.
PR 20-DEC-2001; 2001JP-00403299.
PR 20-DEC-2001; 2001JP-00403300.
PR 27-SEP-2002; 2002JP-00327515.
XX
XX (UYN1-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 569 BP; 137 A; 175 C; 108 G; 148 T; 0 U; 1 Other;
Query Match 3.9%; Score 46.6; DB 9; Length 569;
Best Local Similarity 56.1%; Pred. No. 0.00039;
Matches 88; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Qy 371 CGTCCTGTGCACATTCTTCTGAATGACATCAGGATCCCGCCGTCTCTACTGGGATAAC 430
Db 119 CTTTCTCTCTGGTCTCCAGGAATAATGTGAGGTATGCTTCCAGCTCGAGCAGCAGCAC 178
Qy 431 GGGCAGCGCGGAGACTGACGCTTACGCCAGTACCAAGCCTTCATTTCCGAGG 490
Db 179 CGGAATCTCCGATGACATTCCTCCAAACACAAACCCAGCGCTCTCTGATTGAGAAG 238
Qy 491 CATGACACACACATCGGCAATCCGGTAGACCGGTAAAC 527
Db 239 CATCAACACACATCCCGCTGCGCATAGGCTGTGAC 275
RESULT 24
ACLI17733
ID ACLI17733 standard; DNA; 609 BP.
XX
AC ACLI17733;
XX
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7724.

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Akashinriki).
XX
FN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
XX
PR 20-DEC-2001; 2001JP-00387131.
XX
PR 20-DEC-2001; 2001JP-00403299.
XX
PR 20-DEC-2001; 2001JP-00403300.
XX
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 609 BP; 140 A; 185 C; 126 G; 158 T; 0 U; 0 Other;

Query Match 3.9%; Score 46.6; DB 9; Length 609;
Best Local Similarity 56.1%; Pred. No. 0.0004;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 371 CGTCCCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 430
DB 423 CTTTCTCTCTGCTCTCAGGAATATGTGAGTATGCTTCCAGCTCGAGCAGCGACCAC 482
QY 431 GGGCAGCGCGAGACTGACGCTTTCAGCCAGTACATACCAAAACGCTTCAATTTCCGAAGG 490
DB 483 CGGAATCTCCCGATGACATTGCTCTCCAAACACAAACCCAGCGTCTCTGATTTCAGAAGG 542
QY 491 CATGACCAACACACTGGCATCCGGTAGACCGGTAAAC 527
DB 543 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 579

RESULT 25
ACLI17706
ID ACLI17706 standard; DNA; 615 BP.
XX
AC ACLI17706;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #7697.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; var. (cul.Akashinriki).
XX
FN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
XX
PR 20-DEC-2001; 2001JP-00387131.
XX
PR 20-DEC-2001; 2001JP-00403299.
XX
PR 20-DEC-2001; 2001JP-00403300.
XX
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
DR WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 615 BP; 141 A; 193 C; 118 G; 163 T; 0 U; 0 Other;

Query Match 3.9%; Score 46.6; DB 9; Length 615;
Best Local Similarity 56.1%; Pred. No. 0.00041;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 371 CGTCCCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 430
DB 172 CTTTCTCTCTGCTCTCAGGAATATGTGAGTATGCTTCCAGCTCGAGCAGCGACCAC 231
QY 431 GGGCAGCGCGAGACTGACGCTTTCAGCCAGTACATACCAAAACGCTTCAATTTCCGAAGG 490
DB 232 CGGAATCTCCCGATGACATTGCTCTCCAAACACAAACCCAGCGTCTCTGATTTCAGAAGG 291
QY 491 CATGACCAACACACTGGCATCCGGTAGACCGGTAAAC 527
DB 292 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 328

RESULT 26
ACLI17711
ID ACLI17711 standard; DNA; 619 BP.
XX
AC ACLI17711;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #7702.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; var. (cul.Akashinriki).
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
XX
XX 20-DEC-2001; 2001JP-00387131.
XX
XX 20-DEC-2001; 2001JP-00403299.
XX
XX 20-DEC-2001; 2001JP-00403300.
XX
XX 27-SEP-2002; 2002JP-00327515.
XX
XX (UYNI-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
XX sequences containing them for analysis and identification of barley
XX varieties and production of barley transformants with desired
XX characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
XX barley (Hordeum vulgare) which contain single nucleotide polymorphisms
XX (SNP). The oligonucleotides may be used for analysis of SNPs among barley
XX varieties, identification of particular varieties and genotype-phenotype
XX analysis, isolation of specific genes and creation of new varieties by
XX transformation of barley varieties with them and production of new barley
XX varieties with desired properties. The present sequence represents an
XX oligonucleotide clone DNA sequence featured in the specification. The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 619 BP; 156 A; 168 C; 134 G; 161 T; 0 U; 0 Other;
XX
XX Query Match 3.9%; Score 46.6; DB 9; Length 619;
XX Best Local Similarity 56.1%; Pred. No. 0.00041;
XX Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
XX
XX 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGGATCCCGCGTCTCTCACTGGCGATAAC 430
XX 1 CTTTCTTCTCTGGTCTCAGGAATAATGTTCAGGTATGCTTCCAGTCCGAGCGACAC 60
XX
XX 431 GGGCAGCCGGAGACTGACGTTTCAGCCAGTACATACCAACAGCTTCATTTCCGAAGG 490
XX 61 CGGAATCTCCGATGACATTCCTCCCAACCAACAAACCCAGCGTCTCTGATTCAGAAGG 120
XX
XX 491 CATGACCAACCACTGGCAATCCGGTAGACCGGTAAAC 527
XX 121 CATCAAAACACATCCCCACTGGCATAGGCTGTGAC 157
XX
XX RESULT 27
XX ACL17726
XX ID ACL17726 standard; DNA; 640 BP.
XX
XX ACL17726;
XX
XX 27-OCT-2003 (revised)
XX DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7717.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; var. (cul.Haruna Nijo).
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
XX
XX 20-DEC-2001; 2001JP-00387131.
XX
XX 20-DEC-2001; 2001JP-00403299.
XX
XX 20-DEC-2001; 2001JP-00403300.
XX
XX 27-SEP-2002; 2002JP-00327515.
XX
XX (UYNI-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
XX sequences containing them for analysis and identification of barley
XX varieties and production of barley transformants with desired
XX characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
XX barley (Hordeum vulgare) which contain single nucleotide polymorphisms
XX (SNP). The oligonucleotides may be used for analysis of SNPs among barley
XX varieties, identification of particular varieties and genotype-phenotype
XX analysis, isolation of specific genes and creation of new varieties by
XX transformation of barley varieties with them and production of new barley
XX varieties with desired properties. The present sequence represents an
XX oligonucleotide clone DNA sequence featured in the specification. The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 640 BP; 156 A; 189 C; 133 G; 161 T; 0 U; 1 Other;
XX
XX Query Match 3.9%; Score 46.6; DB 9; Length 640;
XX Best Local Similarity 56.1%; Pred. No. 0.00042;
XX Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
XX
XX 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGGATCCCGCGTCTCTCACTGGCGATAAC 430
XX 484 CTTTCTTCTCTGGTCTCAGGAATAATGTTCAGGTATGCTTCCAGTCCGAGCGACAC 543
XX
XX 431 GGGCAGCCGGAGACTGACGTTTCAGCCAGTACATACCAACAGCTTCATTTCCGAAGG 490
XX 544 CGGAATCTCCGATGACATTCCTCCCAACCAACAAACCCAGCGTCTCTGATTCAGAAGG 603
XX
XX 491 CATGACCAACCACTGGCAATCCGGTAGACCGGTAAAC 527
XX 604 CATCAAAACACATCCCCACTGGCATAGGCTGTGAC 640
XX
XX RESULT 28
XX ACL17732
XX ID ACL17732 standard; DNA; 647 BP.
XX
XX ACL17732;
XX
XX 27-OCT-2003 (revised)
XX DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7723.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; ssp. spontaneum.
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
XX
XX 20-DEC-2001; 2001JP-00387131.
XX
XX 20-DEC-2001; 2001JP-00403299.
XX
XX 20-DEC-2001; 2001JP-00403300.
XX
XX 27-SEP-2002; 2002JP-00327515.
XX
XX (UYNI-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 647 BP; 151 A; 195 C; 133 G; 165 T; 0 U; 3 Other;
SQ
Query Match 3.9%; Score 46.6; DB 9; Length 647;
Best Local Similarity 56.1%; Pred. No. 0.00042;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCCGTCCTCACTGGCGATAAC 430
Db 460 CTTTCTTCTGTCCTCAGGAATATGTCAGGTATGCTTCCAGCTCGAGCGACCAC 519
QY 431 GGCGACCGCGGAGACTGACGCTTCAGCCAGTACCATACCAAAACGCTTCAATTTTCCGAAGG 490
Db 520 CGGAATCCCGATGACATTGCTCTCCACACAAACCAACCCAGCGTCTCTGATTCAAGAAGG 579
QY 491 CATGACCACCACTGGCAATCCGGTAGACCGGTAAAC 527
Db 580 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 616
RESULT 29
ACLI17727
ID ACLI17727 standard; DNA; 651 BP.
XX
XX ACLI17727;
XX
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7718.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; ssp. spontaneum.
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
XX
XX 20-DEC-2001; 2001JP-00387131.
XX
XX 20-DEC-2001; 2001JP-00403299.
XX
XX 20-DEC-2001; 2001JP-00403300.
XX
XX 27-SEP-2002; 2002JP-00327515.
XX
XX (UYNI-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 651 BP; 153 A; 199 C; 134 G; 165 T; 0 U; 0 Other;
SQ
Query Match 3.9%; Score 46.6; DB 9; Length 651;
Best Local Similarity 56.1%; Pred. No. 0.00042;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCCGTCCTCACTGGCGATAAC 430
Db 443 CTTTCTTCTGTCCTCAGGAATATGTCAGGTATGCTTCCAGCTCGAGCGACCAC 502
QY 431 GGCGACCGCGGAGACTGACGCTTCAGCCAGTACCATACCAAAACGCTTCAATTTTCCGAAGG 490
Db 503 CGGAATCCCGATGACATTGCTCTCCACACAAACCAACCCAGCGTCTCTGATTCAAGAAGG 562
QY 491 CATGACCACCACTGGCAATCCGGTAGACCGGTAAAC 527
Db 563 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 599
RESULT 30
ACLI17722
ID ACLI17722 standard; DNA; 669 BP.
XX
XX ACLI17722;
XX
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7713.
```



```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; var. (cul.Haruna Nijo).
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
XX
XX 20-DEC-2001; 2001JP-00387131.
XX
XX 20-DEC-2001; 2001JP-00403299.
XX
XX 20-DEC-2001; 2001JP-00403300.
XX
XX 27-SEP-2002; 2002JP-00327515.
XX
XX (UYNI-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 669 BP; 161 A; 202 C; 135 G; 171 T; 0 U; 0 Other;
SQ
Query Match 3.9%; Score 46.6; DB 9; Length 669;
Best Local Similarity 56.1%; Pred. No. 0.00043;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGGATCCCGCCGTCTCTCACTGGCGATAAC 430
Db 460 CTTTCCTTCTCGTCTCAGGAATAATGTTCAGGTATGCTTCCAGCTCGAGCAGCGACCA 519
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCGATGACATACCAACAGCTTCATTTCCGAGG 490
Db 520 CGGAACCTCCCGATGACATTCCTTCCAAACAACAACACCCAGCGTCTCTGATTCAGAGG 579
QY 491 CATGACCACCACTGGCAATCCGGTAGACCGGTAAC 527
Db 580 CATCACAACACATCCCCACTGGCATAGGCTGTGAC 616
RESULT 31
ACLI17735
ID ACLI17735 standard; DNA; 663 BP.
XX
XX ACLI17735;
AC
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7726.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
XX Hordeum vulgare; var. (cul.Haruna Nijo).
XX
XX WO2003057877-A1.
XX
XX 17-JUL-2003.
XX
XX 16-DEC-2002; 2002WO-IB005403.
XX
XX 20-DEC-2001; 2001JP-00387059.
XX
XX 20-DEC-2001; 2001JP-00387131.
XX
XX 20-DEC-2001; 2001JP-00403299.
XX
XX 20-DEC-2001; 2001JP-00403300.
XX
XX 27-SEP-2002; 2002JP-00327515.
XX
XX (UYNI-) UNIV JAPAN OKAYAMA.
XX
XX Sato K, Takeda K, Kohara Y;
XX WPI; 2003-587127/55.
XX
XX Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
XX Disclosure; SEQ ID XX; 284pp; Japanese.
XX
XX The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 683 BP; 167 A; 203 C; 139 G; 172 T; 0 U; 2 Other;
SQ
Query Match 3.9%; Score 46.6; DB 9; Length 683;
Best Local Similarity 56.1%; Pred. No. 0.00043;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGGATCCCGCCGTCTCTCACTGGCGATAAC 430
Db 473 CTTTCCTTCTCGTCTCAGGAATAATGTTCAGGTATGCTTCCAGCTCGAGCAGCGACCA 532
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCGATGACATACCAACAGCTTCATTTCCGAGG 490
Db 533 CGGAACCTCCCGATGACATTCCTTCCAAACAACAACCCAGCGTCTCTGATTCAGAGG 592
QY 491 CATGACCACCACTGGCAATCCGGTAGACCGGTAAC 527
Db 593 CATCACAACACATCCCCACTGGCATAGGCTGTGAC 629
RESULT 32
ACLI17731
ID ACLI17731 standard; DNA; 693 BP.
XX
XX ACLI17731;
AC
XX 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
XX DNA clone originating in barley containing SNP encoding sequence #7722.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; ssp. spontaneum.
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
XX
PR 20-DEC-2001; 2001JP-00387131.
XX
PR 20-DEC-2001; 2001JP-00403299.
XX
PR 20-DEC-2001; 2001JP-00403300.
XX
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
XX WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 693 BP; 170 A; 204 C; 143 G; 175 T; 0 U; 1 Other;
Query Match 3.9%; Score 46.6; DB 9; Length 693;
Best Local Similarity 56.1%; Pred. No. 0.00044;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTCTTCGTAATGACATCAGGGATCCCGCCGTCCTCACTGGCGATAAC 430
Db 487 CTTTCCTTCCTGTCCTCAGGAATATGTCAGGTATGCTCCAGCTCGAGCAGCACC 546
QY 431 GGGCAGCGCGAGACTGACGCTTCAGCCAGTACCATACCAAAACGCTTCATTTCCGAAGG 490
Db 547 CGGAACTCCCGATGACATTCGCTCCAAACAACAACAAACCCAGCGTCTCTGATTCAAGAAG 606
QY 491 CATGACCCACACTGGCATCCGGTAGACCGGTAAC 527
Db 607 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 643
RESULT 33
ACLI17737
ID ACLI17737 standard; DNA; 713 BP.
XX
AC ACLI17737;
XX
DT 27-OCT-2003 (revised)
DT 17-OCT-2003 (first entry)
XX
DE DNA clone originating in barley containing SNP encoding sequence #7728.
```

```
XX Barley; single nucleotide polymorphism; SNP; genotype-phenotype analysis;
KW gene; ss.
XX
OS Hordeum vulgare; ssp. spontaneum.
XX
PN WO2003057877-A1.
XX
PD 17-JUL-2003.
XX
PF 16-DEC-2002; 2002WO-IB005403.
XX
PR 20-DEC-2001; 2001JP-00387059.
XX
PR 20-DEC-2001; 2001JP-00387131.
XX
PR 20-DEC-2001; 2001JP-00403299.
XX
PR 20-DEC-2001; 2001JP-00403300.
XX
PR 27-SEP-2002; 2002JP-00327515.
XX
PA (UYNI-) UNIV JAPAN OKAYAMA.
XX
PI Sato K, Takeda K, Kohara Y;
XX
XX WPI; 2003-587127/55.
XX
PT Single nucleotide polymorphism sites in barley varieties and DNA
PT sequences containing them for analysis and identification of barley
PT varieties and production of barley transformants with desired
PT characteristics.
XX
PS Disclosure; SEQ ID XX; 284pp; Japanese.
XX
CC The present invention relates to oligonucleotide clones originating in
CC barley (Hordeum vulgare) which contain single nucleotide polymorphisms
CC (SNP). The oligonucleotides may be used for analysis of SNPs among barley
CC varieties, identification of particular varieties and genotype-phenotype
CC analysis, isolation of specific genes and creation of new varieties by
CC transformation of barley varieties with them and production of new barley
CC varieties with desired properties. The present sequence represents an
CC oligonucleotide clone DNA sequence featured in the specification. The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published-pct-sequences. (Updated on 27-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 713 BP; 173 A; 209 C; 148 G; 182 T; 0 U; 1 Other;
Query Match 3.9%; Score 46.6; DB 9; Length 713;
Best Local Similarity 56.1%; Pred. No. 0.00044;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTCTTCGTAATGACATCAGGGATCCCGCCGTCCTCACTGGCGATAAC 430
Db 502 CTTTCCTTCCTGTCCTCAGGAATATGTCAGGTATGCTCCAGCTCGAGCAGCACC 561
QY 431 GGGCAGCGCGAGACTGACGCTTCAGCCAGTACCATACCAAAACGCTTCATTTCCGAAGG 490
Db 562 CGGAACTCCCGATGACATTCGCTCCAAACAACAACAAACCCAGCGTCTCTGATTCAAGAAG 621
QY 491 CATGACCCACACTGGCATCCGGTAGACCGGTAAC 527
Db 622 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 658
RESULT 34
ABD06478
ID ABD06478 standard; DNA; 570 BP.
XX
AC ABD06478;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #5082.
```

KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
XX
OS Pseudomonas aeruginosa.
PN US6551795-B1.
XX
XX 22-APR-2003.
PD
XX 18-FEB-1999; 99US-00252991.
XX PF
XX 18-FEB-1998; 98US-0074788P.
XX PR
XX 27-JUL-1998; 98US-0094190P.
XX PR
XX (GENO-) GENOME THERAPEUTICS CORP.
XX PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX DR
XX P-PSDB; ABO72907.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
PT
XX
XX Disclosure; SEQ ID NO 5082; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX SQ Sequence 570 BP; 105 A; 200 C; 193 G; 72 T; 0 U; 0 Other;
Query Match 3.9%; Score 46.2; DB 11; Length 570;
Best Local Similarity 57.1%; Pred. No. 0.00053;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 374 CCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGCGGATAACGGG 433
Db 2 CCGGGTCACGCCCTGCTGCAATACATCCAGCGGCCCGCGCAGCGAAGCGCCGACCGG 61
QY 434 CAGCCGGAGACTGACGCTTCAGCCAGTACATACCAAGCGCTTCATTTCCGAGGCAT 493
Db 62 CAGCCACAGGCCAGGCGCTTCGAGCATCACCAGGCCATAGGTATCGGTACGCGAGGGGAA 121
QY 494 GACCACACACTGGCAATCCGGTAGAC 520
Db 122 CACCAGCACCAGGCGCCCGCGGTAGGC 148
RESULT 35
ABD06561/c
ID ABD06561 standard; DNA; 714 BP.
XX
XX AC ABD06561;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polynucleotide #5165.
XX

KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
XX
OS Pseudomonas aeruginosa.
PN US6551795-B1.
XX
XX 22-APR-2003.
PD
XX 18-FEB-1999; 99US-00252991.
XX PF
XX 18-FEB-1998; 98US-0074788P.
XX PR
XX 27-JUL-1998; 98US-0094190P.
XX PR
XX (GENO-) GENOME THERAPEUTICS CORP.
XX PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX DR
XX P-PSDB; ABO72990.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
PT
XX
XX Disclosure; SEQ ID NO 5165; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX SQ Sequence 714 BP; 90 A; 234 C; 268 G; 122 T; 0 U; 0 Other;
Query Match 3.9%; Score 46.2; DB 11; Length 714;
Best Local Similarity 57.1%; Pred. No. 0.0006;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 374 CCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCCCGTCTCACTGCGGATAACGGG 433
Db 526 CCGGGTCACGCCCTGCTGCAATACATCCAGCGGCCCGCGCAGCGAAGCGCCGACCGG 467
QY 434 CAGCCGGAGACTGACGCTTCAGCCAGTACATACCAAGCGCTTCATTTCCGAGGCAT 493
Db 466 CAGCCACAGGCCAGGCGCTTCGAGCATCACCAGGCCATAGGTATCGGTACGCGAGGGGAA 407
QY 494 GACCACACACTGGCAATCCGGTAGAC 520
Db 406 CACCAGCACCAGGCGCCCGCGGTAGGC 380
RESULT 36
ABK78820/c
ID ABK78820 standard; DNA; 783 BP.
XX
XX AC ABK78820;
XX
XX 13-AUG-2002 (first entry)
XX
XX Bacillus clausii genomic sequence tag (GST) #1663.
XX

KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
XX physiological provocation; ds.

OS Bacillus clausii.

XX W0200229113-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US031437.

XX 06-OCT-2000; 2000US-00680598.

PR 27-MAR-2001; 2001US-0279526P.

XX (NOVO) NOVOZYMES BIOTECH INC.

PA (NOVO) NOVOZYMES AS.

XX Berka R, Clausen IG;

XX WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.

PS Claim 11; SEQ ID NO 6111; 200pp; English.

XX The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC -up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: the sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 783 BP; 224 A; 150 C; 204 G; 203 T; 0 U; 2 Other;

Query Match 3.9%; Score 46.2; DB 6; Length 783;
Best Local Similarity 59.5%; Pred. No. 0.00063;
Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 354 TCCCCACGGGTAAACAGCGTCCCTGTCCACATTCTTCTGAATGACATCAGGGATCCCGCC 413

Db 633 TTCCCAAGGGGGCATAAATACCAAGTTACTTCTGTTTCAACAATTCGCGGATGCCACCA 574

QY 414 GTCTCACTGCGGATACGGGACGGGACGCTTTCAGCCAGTACCATACCAAAAC 473

Db 573 ATTTTGTGCGATCACAGGTACCCCGCATGCCATCGCTTCAAGCGCCACTAAACCAAG 514

QY 474 GCTTCATTTC 484

Db 513 CTTTCTTTTC 503

RESULT 37

ABD06412

ID ABD06412 standard; DNA; 1251 BP.

XX ABD06412;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polynucleotide #5016.

XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

PR 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

DR P-PSDB; AB072841.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 5016; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC the sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 1251 BP; 224 A; 448 C; 415 G; 164 T; 0 U; 0 Other;

Query Match 3.9%; Score 46.2; DB 11; Length 1251;
Best Local Similarity 57.1%; Pred. No. 0.00083;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 374 CCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCCGTCCTCACTGCGGATAACGGG 433

Db 421 CCGGTCACGCCCTGCTGCAATATACATCCAGCGGCCCGGACGCGGAGCCGCCACCGG 480

QY 434 CAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAAGCGTTTCATTTTCCGAAGGCAT 493

Db 481 CAGGCCACAGGCCAGGGCTTCGAGCATCACAGGCCCATAGGTATCGTCAGCGAGGGGAA 540

QY 494 GACCACCACACTGGCAATCGGTAGAC 520

Db 541 CACCAGCAGCGAGGCCCGCCCGGTAGGC 567

RESULT 38

ABD06505/c

ID ABD06505 standard; DNA; 1584 BP.

Db	926	TTTTTCGTGCCGACAACTTTCAGGTATGCTCCAAATATAACTTCGCAATTACCGGTGCGCCA	867
Qy	441	GAGACTGAGCTTCAGCCAGTACCATACCAACGCTTCAT	480
Db	866	CAGGCATGCTTCGCAATTCGATGCCAAATGCTTCAT	827

RESULT 40

RESULI 40
ACF65374/c
ID ACF65374 standard; DNA; 69727 BP.

AC ACF65374:

DT 20-NOV-2003 (first entry)

DE photorhabdus luminescens nucleotide sequence #27.

Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 detection; food; gene expression; plant; animal; microorganism; toxin;
 antibiotic; biopesticide; virulence factor; disease model; plague;
 whooping cough; gene; ds.

OS Photorhabdus luminescens.

PN WO200294867-A2.

XX
PD
28-NOV-2002.

07-FEB-2002: 2002WO-IB003040.

07-FEB-2001: 2001ER-00001659.

PA (INSP) INST PASTEUR.

FA (INSP) INSI FASIEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.

AA Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;

WPI: 2003-148459/14.

Genomic sequence of *Photobacterium luminescens* and encoded polypeptides, useful e.g. as therapeutic antimicrobials and agricultural pesticides.

PS Claim 1; SEQ ID NO 27; 1205pp: French.

The invention relates to the isolation of genes and their encoded proteins from *Photobacterium luminescens*. The isolated sequences are sources of probes and primers for detecting the genome of *P. luminescens* and related species; to study polymorphisms; for gene analysis and for detection/amplification of the genes. Antibodies (Ab) raised against the polypeptides encoded by the genes are used for detection/identification of *P. luminescens*, e.g. in foods. The genes, proteins, Ab and cells that carry a gene-containing vector are used to select compounds that modulate, regulate, induce or inhibit expression of the genes in plants, animals or microorganisms other than *P. luminescens* and are able to alter response or sensitivity to toxins and antibiotics produced by *P.*

SQ Sequence 69727 BP; 20213 A; 13239 C; 14632 G; 21638 T; 0 U; 5 Other;

Query Match	3.9%	Score 45.6;	DB 10;	Length 69727;
Best Local Similarity	53.6%;	Pred. No. 0.014;		
Matches 118;	Conservative	0;	Mismatches 99;	Indels 33;
				Gaps 1;

Qy	264	TCATATCTTCTCTGGCGCTGCTGCCATCATCCGGAAGCGTTCCGGTGGGGATAAAAA	323
Db	31511	TCAATCCGTTGACGAGCTGCTTTTCCCATTTTTTGGCCCCCTGTCGGTGGAGACAATAA	31452
Qy	324	TCGGCGAGTGGCGCGGTCCATGACAGACATCCCCACGGGTAAACAGCGTCCCTGTGCACA	383
Db	31451	AAATTAAACGCGATCGGCAATTCCTGAAGCATCGCTGGCGGTACTAAATACCCGAATTA	31392
Qy	384	TTCTTCTGAATGACA---TCAGGGATCCGCGCGTCTCACTGGCGATAAACGGGCACGCGC	440
Db	31391	TTTTCGTTCGGGACAACTTCAGGTATGCTCCAATATAACTCGCAATTACCGTCGGCCA	31332
Qy	441	GAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCAT	480
Db	31331	CAGGCATCGCTTCGGCAATTGTGATCCAAATGCTTCAT	31292

Search completed: June 4, 2005, 12:32:11
Job time : 732 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 11:53:41 ; Search time 236 Seconds
(without alignments)
8188.317 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgcagagatggaagaaaa.....ttttactttttctctgcag 1181

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.2	4.8	4403765	3 US-09-103-840A-2	Sequence 2, Appli
2	57.2	4.8	4411529	3 US-09-103-840A-1	Sequence 1, Appli
3	50	4.2	948	4 US-09-902-540-8665	Sequence 8665, Ap
4	50	4.2	8544	4 US-09-902-540-917	Sequence 917, App
5	49.6	4.2	601	4 US-09-949-016-50324	Sequence 50324, A
6	49.6	4.2	601	4 US-09-949-016-50325	Sequence 50325, A
7	49.6	4.2	198942	4 US-09-949-016-13209	Sequence 13209, A
8	46.2	3.9	570	4 US-09-252-991A-5082	Sequence 5082, Ap
9	46.2	3.9	714	4 US-09-252-991A-5165	Sequence 5165, Ap
10	46.2	3.9	1251	4 US-09-252-991A-5016	Sequence 5016, Ap
11	46.2	3.9	1584	4 US-09-252-991A-5109	Sequence 5109, Ap
12	41	3.5	1146	4 US-09-902-540-8337	Sequence 8337, Ap
13	41	3.5	8578	4 US-09-902-540-871	Sequence 871, App
14	40.6	3.4	129908	4 US-09-585-858-1	Sequence 1, Appli
15	40.6	3.4	129908	4 US-10-270-878-1	Sequence 1, Appli
16	39.8	3.4	4403765	3 US-09-103-840A-2	Sequence 2, Appli
17	39.8	3.4	4411529	3 US-09-103-840A-1	Sequence 1, Appli
18	37	3.1	855	4 US-09-252-991A-16409	Sequence 16409, A
19	37	3.1	1155	4 US-09-252-991A-16524	Sequence 16524, A
20	37	3.1	1407	4 US-09-252-991A-16074	Sequence 16074, A
21	36.6	3.1	654	4 US-09-825-561A-7	Sequence 7, Appli
22	36.6	3.1	1614	4 US-09-404-641-4	Sequence 4, Appli
23	36.6	3.1	1614	4 US-10-414-186-4	Sequence 4, Appli
24	36.6	3.1	1614	4 US-10-243-072-4	Sequence 66, Appl
25	36.2	3.1	741	4 US-09-825-561A-66	Sequence 16182, A
26	36.2	3.1	1290	4 US-09-252-991A-16182	Sequence 50326, A
27	35	3.0	601	4 US-09-949-016-50326	

C 28 35 3.0 1110 4 US-09-543-681A-3342 Sequence 3342, Ap
C 29 34.8 2.9 42988 4 US-08-311-731A-128 Sequence 128, App
C 30 34.6 2.9 355 4 US-09-513-999C-26063 Sequence 26063, A
C 31 34.2 2.9 2135 4 US-08-933-711B-17 Sequence 17, Appl
C 32 34.2 2.9 24740 4 US-09-949-016-13528 Sequence 13528, A
C 33 34.2 2.9 74962 4 US-09-685-853A-3 Sequence 3, Appli
C 34 34 2.9 1842 4 US-09-489-039A-6050 Sequence 6050, Ap
C 35 33.8 2.9 1170 4 US-09-902-540-8333 Sequence 8353, Ap
C 36 33.8 2.9 1647 6 5405943-3 Patent No. 5405943
C 37 33.8 2.9 1647 6 5405943-3 Patent No. 5405943
C 38 33.8 2.9 1650 4 US-09-252-991A-8097 Sequence 8097, Ap
C 39 33.8 2.9 2145 4 US-09-543-681A-8205 Sequence 8205, Ap
C 40 33.8 2.9 8056 4 US-09-902-540-874 Sequence 874, App
C 41 33.8 2.9 144362 4 US-09-949-016-16066 Sequence 16066, A
C 42 33.6 2.8 546 4 US-09-902-540-7296 Sequence 7296, Ap
C 43 33.6 2.8 1074 4 US-09-902-540-7297 Sequence 7297, Ap
C 44 33.6 2.8 1158 4 US-09-543-681A-3266 Sequence 3266, Ap
C 45 33.6 2.8 4127 4 US-09-902-540-2329 Sequence 2329, Ap

ALIGNMENTS

RESULT 1

US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 4.8%; Score 57.2; DB 3; Length 4403765;
Best Local Similarity 54.2%; Pred. No. 4.6e-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 272 TTCTCTGGCGCTGGCTGCATCCGGAAGCGTTCGGTCCGGGATAAAAAATCGCGCAG 331
Db 2447373 TTCTCGCGCGCGCGCCCATCGGACGCGCCCATCGCATCGCACTCGGCCAC 2447432
QY 332 TGGCGCGCTCCATGCAGACACATCCCCACGGGTAACAGCGTCCCTGTCACTTCCTG 391
Db 2447433 GGCGTCGGCCACCGCGTCCACCGACCTACGTCGACCACTAGCCAGTCTGTGTGCTG 2447492
QY 392 ATGACATCAGGATCCCGCGCTCTCACTGGGGATAACGGGCACGGCGAGACTGACGC 451
Db 2447493 CACCGTTTCGGCGCTCCCGCAGAAATTCGGGGGATACCGGCACGCGCGGAGGC 2447552
QY 452 TTCAGCAGTACCATACCAACGCTTCATTTTCC 485
Db 2447553 TTCAGGAACACGATGCCCAAGCCCTCGAGTCC 2447586

RESULT 2

US-09-103-840A-1
; Sequence 1, Application US/09103840A

Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 4.8%; Score 57.2; DB 3; Length 4411529;
Best Local Similarity 54.2%; Pred. No. 4.6e-06;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 272 TTCTCTGGCGCTGGCTGCATCATCCGGAAGCGTTCCGTCGGGATAAAAAATCGCGCAG 331
Db 2450072 TTCTCGCGCGCGCGCCCATCGCGACGCGCCCGATCCGATCGATCAGCAACTCGGCCAC 2450131

QY 332 TGGCGCGGTCATGCAGACATCCCGCAGCGGTACAGCGGTCCCTGTACATCTTCTCTG 391
Db 2450132 GCGGTGCGGCCACCGCGTCCACCGACCTACCGTCGACCATCTAGGCCAGTCTTGTGTGCTG 2450191

QY 392 AATGACATCAGGATCCCGCGCTCTCCTGCGGATAAAGCGCACGCGGAGACTGACGC 451
Db 2450192 CACGGTTTCGGCGCTCCGCGAGAAATGCGCGGATACCGGACGCGCGCGGAGGC 2450251

QY 452 TTCAGCAGTACCATACCAACGCTTCATTTCC 485
Db 2450252 TTCAGGAACATGATGCCAAGCCCTCGAGCTCC 2450285

RESULT 3
US-09-902-540-8665/c
; Sequence 8665, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8665
; LENGTH: 948
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8665

Query Match 4.2%; Score 50; DB 4; Length 948;
Best Local Similarity 49.6%; Pred. No. 4.7e-06;
Matches 128; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 250 TATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCCGGAAGCGTTCCG 309
Db 880 TCTGGAAGCGTTCCAGGACCTGTGCGCGCGGTGTGAGAAGCCCTGCCAACGCTCCG 821

QY 310 GTCCGATAAAAAATCGCGAGTGGCGGCTTCCATGCAGACACATCCCCACCGGTAACA 369

Db 820 CGTCTCTCGACCAATGTGAGCAGCTGCGGCGCCATGGCCGAGCGTCACCCAGCGGCGCCA 761

QY 370 GCGTCCCTGTGCACATTTCTGAATGACATCAGGATCCGCGCTCTCACTCGCGATAA 429

Db 760 GGAAGCCCGTCTCCCATGGTGACGAGCTCCGGGATTCGCGCCAGGTCTGCGCACCA 701

QY 430 CGGCGACGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAACGCTTCAATTTCCGAAG 489

Db 700 CGGGAATGCGCAGCTGAGCGCTTCCAGCGCGGAGCGCCGAAAGCTCTCTCTGCTCGCTGG 641

QY 490 GCATGACCAACCACTGG 507

Db 640 GGAGCAGGAAGACGTCGG 623

RESULT 4
US-09-902-540-917
; Sequence 917, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 917
; LENGTH: 8544
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; LOCATION: (1)..(8544)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-917

Query Match 4.2%; Score 50; DB 4; Length 8544;
Best Local Similarity 49.6%; Pred. No. 2e-05;
Matches 128; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 250 TATCAACCGGTACTCAATATCTTCTGCGCTGGCTGCATCATCCGGAAGCGTTCCG 309

Db 7663 TCTGGAAGCGTTCCAGGACCTGTGCGCGCGGTGTGAGAAGCCCTGCCAACGCTCCG 7722

QY 310 GTCCGATAAAAAATCGCGAGTGGCGGCTTCCATGCAGACACATCCCCACCGGTAACA 369

Db 7723 CGTCTCTCGACCAATGTGAGCAGCTGCGCGCCATGCGCGGAGTCAACCAGCGCGCCA 7782

QY 370 GCGTCCCTGTGCACATTTCTGAATGACATCAGGATCCGCGCTCTCACTCGCGATAA 429

Db 7783 GGAAGCCCGTCTCCCATGGTGACGAGCTCCGGGATTCGCGCCAGGTCTGCGCACCA 7842

QY 430 CGGCGACGCGGAGACTGACGCTTTCAGCCAGTACCATACCAACGCTTCAATTTCCGAAG 489

Db 7843 CGGGAATGCGCAGCTGAGCGCTTCCAGCGCGGAGCGCCGAAAGCTCTCTCTGCTCGCTGG 7902

QY 490 GCATGACCAACCACTGG 507

Db 7903 GGAGCAGGAAGACGTCGG 7920

RESULT 5
US-09-949-016-50324/c
; Sequence 50324, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED


```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50324
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-50324

Query Match 4.2%; Score 49.6; DB 4; Length 601;
Best Local Similarity 52.4%; Pred. No. 4.7e-06;
Matches 109; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 74 TCTGTCGGTATTTAAATGCAATGACCGTCCCGGTATTTAAACAATGTGATAAATTACTC 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 TCTGTTGGAGCTTAGTGCCATGCACTCAGTCTTTGGTGACATCAGAAATCACCTAGTC 191
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 134 CGTTACCGGAAAACCGCTGAAACAAATTCGGGCTGAAAAGAGGATCGCCGGTTATCTGTT 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 190 CCTGTCCAAACACACTGAATCAGAAATTCGAGCGATGAGACTCTAGTGAATGTGAAGTT 131
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 194 GCATTTCCCTAGCTGATGACGACGACACAAATGATCTGTCGGTCTGTTAATATC 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 130 GTATTTCCCTGTTGGTCTCATTTCTCTGTAATGATGTCAGGCATTCTGTCATTAGC 71
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 254 AAACCGGTACTCAATATCTTCTCTGGCG 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 ATACCACTCATTTGATTTTCTCTGGAG 43
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 6
US-09-949-016-50325/c
; Sequence 50325, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50325
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-50325

Query Match 4.2%; Score 49.6; DB 4; Length 601;
Best Local Similarity 52.4%; Pred. No. 4.7e-06;
Matches 109; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 74 TCTGTCGGTATTTAAATGCAATGACCGTCCCGGTATTTAAACAATGTGATAAATTACTC 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 588 TCTGTTGGAGCTTAGTGCCATGCACTCAGTCTTTGGTGACATCAGAAATCACCTAGTC 529
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 134 CGTTACCGGAAAACCGCTGAAACAAATTCGGGCTGAAAAGAGGATCGCCGGTTATCTGTT 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 528 CCTGTCCTCCAAACACACTGAATCAGAAATTCGACCGATGAGACTCTAGTGAATGTGAAGTT 469
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 194 GCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTCGGTCTGTTAAATATC 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 468 GTATTTCCCTTGTGCTCTCATTTCTCTGTTGACTAATGATGTCAGGCATTCTGTCATTAGC 409
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 254 AAACCGGTACTCAATATCTTCTCTGGCG 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 408 ATACCACTCATTTGTTGTTATTTTCTCTGGAG 381
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
US-09-949-016-13209
; Sequence 13209, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13209
; LENGTH: 198942
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(198942)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13209

Query Match 4.2%; Score 49.6; DB 4; Length 198942;
Best Local Similarity 52.4%; Pred. No. 0.00023;
Matches 109; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 74 TCTGTCGGTATTTAAATGCAATGACCGTCCCGGTATTTAAACAATGTGATAAATTACTC 133
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25279 TCTGTTGGAGCTTAGTGCCATGCACTCAGTCTTTGGTGACATCAGAAATCACCTAGTC 25338
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 134 CGTTACCGGAAAACCGCTGAAACAAATTCGGGCTGAAAAGAGGATCGCCGGTTATCTGTT 193
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25339 CCTGTCCTCCAAACACACTGAATCAGAAATTCGACCGATGAGACTCTAGTGAATGTGAAGTT 25398
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 194 GCATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTCGGTCTGTTAATATC 253
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25399 GTATTTCCCTTGTGCTCATTTCTCTGTTGACTAATGATGTCAGGCATTCTGTCATTAGC 25458
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 254 AAACCGGTACTCAATATCTTCTCTGGCG 281
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25459 ATACCACTCATTTGTTGTTATTTTCTCTGGAG 25486
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-09-252-991A-5082
; Sequence 5082, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5082
; LENGTH: 570
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5082

Query Match 3.9%; Score 46.2; DB 4; Length 570;
Best Local Similarity 57.1%; Pred. No. 6.6e-05;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 374 CCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCGTCTCACTGGCGATAACGGG 433
Db 2 CCGGTTCACGCCCTGCTGCAATACATCCAGCGGCCCGGACGCGAAGCCGCCACCGG 61

QY 434 CAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAAGCGTTTCATTTTCCGAAGGCAT 493
Db 62 CAGGCCACAGCCAGGCGTTTCGAGCATCACCAGGCCATAGGTATCGGTACGCGAGGGAA 121

QY 494 GACCACCACTGGCAATCCGGTAGAC 520
Db 122 CACCAGCACCGAGCGCGCCGGTAGGC 148

RESULT 9

US-09-252-991A-5165/c
; Sequence 5165, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5165
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5165

Query Match 3.9%; Score 46.2; DB 4; Length 714;
Best Local Similarity 57.1%; Pred. No. 7.7e-05;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 374 CCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCGTCTCACTGGCGATAACGGG 433
Db 526 CCGGTTCACGCCCTGCTGCAATACATCCAGCGGCCCGGACGCGAAGCCGCCACCGG 467

QY 434 CAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAAGCGTTTCATTTTCCGAAGGCAT 493
Db 466 CAGGCCACAGGCGAGGCGTTTCGAGCATCACCAGGCCATAGGTATCGGTACGCGAGGGAA 407

QY 494 GACCACCACTGGCAATCCGGTAGAC 520
Db 406 CACCAGCACCGAGCGCGCCGGTAGGC 380

RESULT 10

US-09-252-991A-5016
; Sequence 5016, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5016
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5016

Query Match 3.9%; Score 46.2; DB 4; Length 1251;
Best Local Similarity 57.1%; Pred. No. 0.00011;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 374 CCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCGTCTCACTGGCGATAACGGG 433
Db 421 CCGGTTCACGCCCTGCTGCAATACATCCAGCGGCCCGGACGCGAAGCCGCCACCGG 480

QY 434 CAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAAGCGTTTCATTTTCCGAAGGCAT 493
Db 481 CAGGCCACAGGCGAGGCGTTTCGAGCATCACCAGGCCATAGGTATCGGTACGCGAGGGAA 540

QY 494 GACCACCACTGGCAATCCGGTAGAC 520
Db 541 CACCAGCACCGAGCGCGCCGGTAGGC 567

RESULT 11

US-09-252-991A-5109/c
; Sequence 5109, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 5109
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5109

Query Match 3.9%; Score 46.2; DB 4; Length 1584;
Best Local Similarity 57.1%; Pred. No. 0.00013;
Matches 84; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 374 CCTGTGCACATTCTTCTGAATGACATCAGGGATCCCGCGTCTCACTGGCGATAACGGG 433
Db 1389 CCGGTTCACGCCCTGCTGCAATACATCCAGCGGCCCGGACGCGAAGCCGCCACCGG 1330

QY 434 CAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAAGCGTTTCATTTTCCGAAGGCAT 493
Db 1329 CAGGCCACAGGCGAGGCGTTTCGAGCATCACCAGGCCATAGGTATCGGTACGCGAGGGAA 1270

QY 494 GACCACCACTGGCAATCCGGTAGAC 520
Db 1269 CACCAGCACCGAGCGCGCCGGTAGGC 1243

RESULT 12

Db 10181 ATCTGCAATACCAACCCGCTCATAGTAGCAACCAACCGGGGCTCCACACAGCGACGCCTCAAG 10122
QY 458 CAGTACCATACCAACCGCTTCATTTCCGAAGGATGACACCAACCACTGGC 508
Db 10121 CGAAGGATCCCGAACCCCTCATTAGATGAAGCTGCAGCACCACATCGGC 10071

RESULT 16

US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 3.4%; Score 39.8; DB 3; Length 4403765;
Best Local Similarity 50.8%; Pred. No. 4;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 333 GCGCCGGTCCATGCAGACATCCCCACGGGTAACAGCGTCCCTGTGCATTTCTTGA 392
Db 577989 GCGTCGGCCCATGACCGACCTGTGCCGGACACAGGGTGCGGTGATCCCGTCGCGC 577930
QY 393 ATGACATCAGGGATCCGCGCGTCTCACTGGCGATACCGGGCACCGCGGAGACTGACGT 452
Db 577929 ACCGCGACGGCAGCGCGCCACCGCGCGCCACACCGGTGTGCGCACGCTTGGGCC 577870
QY 453 TCAGCCAGTACCATACCAACGCTTCATTTCCGAAGGATGACCAACCACTGGCAATC 512
Db 577869 TCACAGCAACGAGCGCGCAAGCACTCGGAGTAGCTCGGACCGCAACCAAGGTCCGCGGCC 577810
QY 513 CGGTAGA 519
Db 577809 CGAAACA 577803

RESULT 17

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 3.4%; Score 39.8; DB 3; Length 4411529;
Best Local Similarity 50.8%; Pred. No. 4;
Matches 95; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 333 GCGCCGGTCCATGCAGACATCCCCACGGGTAACAGCGTCCCTGTGCATTTCTTGA 392
Db 576547 GCGTCGGCCCATGACCGACCTCGTCCCGGACACCAAGGGTGCGGTGATCCCGTCGCGC 576488
QY 393 ATGACATCAGGGATCCGCGCGTCTCACTGGCGATACCGGGCACCGCGGAGACTGACGT 452
Db 576487 ACCGCGACGGCAGCGCGCCACCGCGCGCCACCAACCGGTGTGCGCACGCTTGGGCC 576428
QY 453 TCAGCCAGTACCATACCAACGCTTCATTTCCGAAGGATGACCAACCACTGGCAATC 512
Db 576427 TCACAGCAACCAAGCGCGCAACGACTCGGAGTAGCTCGGCACCGCAACCAAGGTCCGCGGCC 576368
QY 513 CGGTAGA 519
Db 576367 CGAAACA 576361

RESULT 18

US-09-252-991A-16409/c
; Sequence 16409, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16409
; LENGTH: 855
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16409

Query Match 3.1%; Score 37; DB 4; Length 855;
Best Local Similarity 47.3%; Pred. No. 0.12;
Matches 149; Conservative 0; Mismatches 160; Indels 6; Gaps 1;

QY 324 TCGCGCAGTGCAGGTCATGCAGACATCCCCACGGGTAACAGCGTCCCTGTGCACA 383
Db 840 TGGGTGAGGCTCGGCGCAGGCGCGTCTCGTCGCCACAGCGGGAAGAGATGCGGACCCCT 781
QY 384 TTCTTCTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATACCGGCGACCGCGAG 443
Db 780 TCC-----ACCACTTCGCGGGCGCGCCACAGCGGTGGCGATCAGCGGTACACCGGCA 727
QY 444 ACTGACGTTTCAGCCAGTACCATACCAACCGTTTCAATTTTCGAAAGGATGACCAACACA 503
Db 726 ACCATGGCTTCAGCAGGACCATCGCGAACCGGCTCGTGGTTCGGAACCTCAGGGGGAAGCG 667
QY 504 CTGGCAATCCGTAGACCGGTAACGCTGGGGAAGGCGACCTGCCATTAAACATCTCCG 563
Db 666 TCGAACGCTTGAAGTAGCGCGCTCTTTCAGCTGCGCGGAGAACACGACGCGCTCG 607
QY 564 CTCATTCCAGGTTTCTGTGTGCTGACGACGAGTCTTCGTTATTTTCACCGCCCGCGC 623
Db 606 CCGATACCCAGTTTCGAGGCCAGGTCTTTGAGCTGCTCTTCCAGGGCGCGCTGCGGAGG 547
QY 624 CCCACCAAGCCAG 638

Query Match 3.1%; Score 36.2; DB 4; Length 1290;
Best Local Similarity 54.9%; Pred. No. 0.3;
Matches 96; Conservative 0; Mismatches 73; Indels 6; Gaps 1;
QY 324 TCGCGAGTGGCGCGTCCATGACAGACATCCCGGTAACGGGTACAGCGTCCCTGTGCACA 393
Db 1099 TGGGTGAGGCGCTCGCGCGAGGCGGTCTCGTCCCGGAGGAGGATGCCGACCCCT 1158
QY 384 TTCTTCTGAATGACATCAGGATCCGCGCGTCTCACTGGCGATAGCGGCAACGGCGGAG 443
Db 1159 TCC-----ACCACTTCGGGGCGCGCCACAGCGGTGGCGATCAGCGGTACACGGCA 1212
QY 444 ACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTTCGAAGGCATGACCA 498
Db 1213 ACCATGCGCTCCAGCAGGACCATGCCAGCGGCTCGTGGTCCGAATCTCAGGGCGA 1267

RESULT 27
US-09-949-016-50326/c
; Sequence 50326, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 50326
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-50326

Query Match 3.0%; Score 35; DB 4; Length 601;
Best Local Similarity 61.5%; Pred. No. 0.46;
Matches 56; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 191 GTTGCATTTCCCTTAGCTGACTAGCCAGACACAATGATGTGCGGTCTGTGTAAT 250
Db 596 GTTGTATTTCCCTTGTGGTCTCATTTCTCTGTGACTAATGATGTGCGGCAATCTGTGCA 537
QY 251 ATCAACCGGTACTCAATATCTTCTGGCG 281
Db 536 AGCATACAGTCAATTTGATTTCTCTGGAG 506

RESULT 28
US-09-543-681A-3342/c
; Sequence 3342, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 3342
; LENGTH: 1110
; TYPE: DNA

; ORGANISM: Proteus mirabilis
US-09-543-681A-3342
Query Match 3.0%; Score 35; DB 4; Length 1110;
Best Local Similarity 57.9%; Pred. No. 0.69;
Matches 62; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
QY 396 ACATCAGGATCCCGCCGCTCTCACTGGCGATAACGGGCACGCGGAGACTGACGGTTCA 455
Db 914 ACTTCAGGAATACCGCGGATAGCTGGCAATCACCGGTTTCCACATGCATGCTTCT 855
QY 456 GCCAGTACCATCAACACGCTTCATTTTCGAAGGCATGACCAACCAC 502
Db 854 GCAATGGTGATCCCAATGCTTCATCGCGGATACTCGGGAACCC 808

RESULT 29
US-08-311-731A-128
; Sequence 128, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 128:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42988 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-128

Query Match 2.9%; Score 34.8; DB 4; Length 42988;
Best Local Similarity 48.1%; Pred. No. 9.3;
Matches 99; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
QY 317 TAAAAAATCGCGCAGTGGCGGTCATGCAACATCCCCCAGGGGTAAACGGGTCCC 376
Db 39694 TAACAGCTGATCGAGCGGTCGCGCAATGACCGACATATATGTCGAACACCAAGTGCC 39753
QY 377 TGTCAATTTCTGAAATGACATCAGGATCCCGCGCTCTCACTGGCGATTAACGGGCAC 436

Db 39754 GGTGACCCCGTCCGACCGCTACCGCAAGCGCGCGCGCTACCGCGAGT 39813
Qy 437 GCCGGAGACTGACGCTTACGCCAGTACCATAACAAAGCGTTTCATTTTCCGAAGGCATGAC 496
Db 39814 GCCACAGCGCTGCGCCTCGACGCGGACCGAGCCCAACAGCACTCTGAATAACTCGGCACCGC 39873
Qy 497 CACCACACTGGCAATCCGGTAGACCG 522
Db 39874 AACTAGATCGCGCGCTTGGGAAGACTG 39899

RESULT 30
US-09-513-999C-26063/c
; Sequence 26063, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 26063
; LENGTH: 355
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 70
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 126
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 175
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 207
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 208
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 219
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 227
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 246
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 273
; OTHER INFORMATION: y=c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 274
; OTHER INFORMATION: y=c or t
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: 286
; OTHER INFORMATION: b=c or g or t
US-09-513-999C-26063
Query Match 2.9%; Score 34.6; DB 4; Length 355;
Best Local Similarity 48.4%; Pred. No. 0.44;
Matches 77; Conservative 6; Mismatches 76; Indels 0; Gaps 0;
Qy 139 CCGGAAACCGCTGAAACAAAATTCGGGCTGAAAGAGGATCCGCCGTTTCTCTTGCATT 198
Db 329 CTGAATAACAAGAGCTCTCAATTTAGGGGGAAGAGCCTCCCAVACTGACATCTARRAAT 270
Qy 199 TCCCTTAGCTGACTAGCCAGAGACACATGATCTGCGCTTCTGTTAATATCAAAACC 258
Db 269 TCCTCTTAGAGACACAGTTACTKCCAAACAAATCTTTTCCTGTTTAATGAATTAACC 210
Qy 259 GGTACTCAATATCTTCTCTGGGCTGGCTGCCATCATCC 297
Db 209 CSRACCTGGTCATATCTTTGAGAGACTGCCTTCATATAC 171

RESULT 31
US-08-933-711B-17/c
; Sequence 17, Application US/08933711B
; Patent No. 6514724
; GENERAL INFORMATION:
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Chuang, Pao-Tien
; TITLE OF INVENTION: HEDGEHOG INTERACTING PROTEINS AND USES RELATED THERETO
; FILE REFERENCE: HUV-024.01
; CURRENT APPLICATION NUMBER: US/08/933,711B
; CURRENT FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/026,155
; PRIOR FILING DATE: 1996-09-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 17
; LENGTH: 2135
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Consensus
; OTHER INFORMATION: sequence
; NAME/KEY: MOD_RES
; LOCATION: (1)..(2135)
; OTHER INFORMATION: "n" base may be a, t, c or g
US-08-933-711B-17

Query Match 2.9%; Score 34.2; DB 4; Length 2135;
Best Local Similarity 29.1%; Pred. No. 2;
Matches 41; Conservative 48; Mismatches 52; Indels 0; Gaps 0;
Qy 254 AAACCGGTACTCAATATCTTCTGGCTGGCTGCATCATCCGGAAGCGTTCCGCTCG 313
Db 798 RWAAGYTCYTTKADYAKTYMYKTCKWGDSTDARDATYWCACRWADCTTCYTYTC 739
Qy 314 GGATAAAAATCGCGAGTGGCGGCTCCATGCAGACACATCCCCACGGGTAAACGCGT 373
Db 738 HARWATRAARAGVCGHTGVGAKCCATCYCCRCWRTGVASVRCDCYNACRGYTGKXAG 679
Qy 374 CCTGTCAATCTTCTGAAT 394
Db 678 HCCRCTMBRAYVTCTGRAY 658

RESULT 32
US-09-949-016-13528
; Sequence 13528, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13528
; LENGTH: 24740
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13528

Query Match 2.9%; Score 34.2; DB 4; Length 24740;
Best Local Similarity 52.4%; Pred. No. 10;
Matches 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 56 AGCTGAAAGTAGCCCTGTTCTGTCGGTATTTAAATGCATTGACCGTCCCGCTATTTTAAA 115
DB 11455 AGCTGAGATCATGCCATTGCATCTCCAGCTGGGCAACAAGAGTGAACCTCATCTCAAAA 11514
QY 116 CAATGTGATAAATTACTTCCGTATACCGGAAACCGCTGAACAAAATTCGGGCTGAAAAGAG 175
DB 11515 AAAAAAGAAAAGAAATGTGTAANTACACATTTGGATACTATTTCAGCCTTAAAAAAG 11574
QY 176 GATCCGCGCTATCTGTTCGATT 198
DB 11575 GAAACTCTGTCTATTGTGACAA 11597

RESULT 33
US-09-685-853A-3
; Sequence 3, Application US/09685853A
; Patent No. 6479270
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; FILE REFERENCE: CL000871
; CURRENT APPLICATION NUMBER: US/09/685,853A
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/182,194
; PRIOR FILING DATE: 2000-02-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 74962
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(74962)
; OTHER INFORMATION: n = A, T, C or G
US-09-685-853A-3

Query Match 2.9%; Score 34.2; DB 4; Length 74962;
Best Local Similarity 52.4%; Pred. No. 22;
Matches 75; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 56 AGCTGAAAGTAGCCCTGTTCTGTCGGTATTTAAATGCATTGACCGTCCCGCTATTTTAAA 115
DB 5571 AGCTGAGATCATGCCATTGCATCTCCAGCTGGGCAACAAGAGTGAACCTCATCTCAAAA 5630
QY 116 CAATGTGATAAATTACTTCCGTATACCGGAAACCGCTGAACAAAATTCGGGCTGAAAAGAG 175
DB 5631 AAAAAAGAAAAGAAATGTGTAANTACACATTTGGATACTATTTCAGCCTTAAAAAAG 5690

QY 176 GATCCGCGCTATCTGTTCGATT 198
DB 5691 GAAACTCTGTCTATTGTGACAA 5713
RESULT 34
US-09-489-039A-6050/c
; Sequence 6050, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO: 6050
; LENGTH: 1842
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6050

Query Match 2.9%; Score 34; DB 4; Length 1842;
Best Local Similarity 45.3%; Pred. No. 2.1;
Matches 124; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 350 CACATCCCCACGGGTAAACAGGTCCCTGTCTCATCTTCTTGAATGACATCAGGATCCC 409
DB 1272 CCGGTTCGCGCCAGCCAGCAGACCAACAGCGAGGTGAGCACCCTCCGCGCCATCGC 1213
QY 410 GCCCGTCTCACTGGGATTAACGGGCACGCGGAGACTGACGCTTACGCCAGTACCATA 469
DB 1212 CGCCAGGCAATCGCGCTCAGGTTCTGCGCGAAGAGGACCATCGCCACGACGCGCGCT 1153
QY 470 AAACGCTTCATTTTTCGAAAGGATGACACCACTGGCAATCCGCTAGACCGGTAAACGC 529
DB 1152 CCAGGCGCGGTGTTAAAGCCCATCAGTCCGGCTCCGAGCGGTACGGGTACGCGA 1093
QY 530 TGGGAAAAGGACCTGCGCATTAACACATCTCCGCTCATCCAGGTGTTCTGTCTGTG 589
DB 1092 CTGGAATATGGCGCCCTGACGCGGCGCGCGATGAGCAGCGCCATTAACACCCG 1033
QY 590 AGCGACGCTGCTGCTATTTCTTACGCGCGCGG 623
DB 1032 CGGACGCGCCCATTCGTTCCACCACCATCGCGAG 999

RESULT 35
US-09-902-540-8353/c
; Sequence 8353, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO: 8353
; LENGTH: 1170
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8353

Query Match 2.9%; Score 33.8; DB 4; Length 1170;

	Best Local Similarity	50.3%; Pred. No. 1.8;	Matches	83; Conservative	0; Mismatches	82; Indels	0; Gaps	0;
Qy	359	CACGGGTAACAAGCGTCCCTGTCACATTTCTTGAAATGACATCAGGAGTCGCCCGCTGC	418					
Db	1038	CACGGGACCAGCACAGGCGCGTCTTCTCGTGTGCGACGATTTCGACGGCGCTCCGGAGCG	.979					
Qy	419	ACTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGCCAGTAGTACCATACCAAACGCTTC	478					
Db	.978	CGTTGCCACACAGCGGTAGACCGGAGGCATGGCTTCGATGCCGTGTAGCCNAAGGCTC	919					
Qy	479	ATTTTCGAAGCGCATGACCAACCACTGGGCAATCCGTFPAGACCGG	523					
Db	918	GTAGAGCGAAGCCACCGACGCGCCAGATCCGCGALTTCGATGGAGCAG	874					

RESULT 36
5405943-3/c
; Patent NO. 5405943
; APPLICANT: COMINGS, DAVID E.
; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED
; BEHAVIORS
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION NUMBER: US/07/562,596
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA
; APPLICATION NUMBER: 125,577
; FILING DATE: 23-NOV-1987
; APPLICATION NUMBER: 271,653
; FILING DATE: 16-NOV-1988
; APPLICATION NUMBER: 410,831
; FILING DATE: 22-SEP-1989
; SEQ ID NO:3
; LENGTH: 1647
5405943-3

	Query Match	2.9%;	Score 33.8;	DB 6;	Length 1647;
	Best Local Similarity	53.4%;	Pred. No. 2.3;		
	Matches	71;	Conservative	0;	Mismatches 62; Indels 0; Gaps 0;
Qy	1013	ATTACTGTCATGGCCACTATGGCAGATGACCAGATCAGGTTTAAATTTCCCGGATAATCCG	1072		
Db	345	ATCTCTGACATGGCCATCTCTGAAAGATCTCTCGAACAGAAATCCAACTCCCGAGGAGTTTG	286		
Qy	1073	TCGAAGTCTCAGGATGCAAGGAAAGGTGCAAGGCTGTTCTCGAAAGGAATAAAAGTGACATC	1132		
Db	285	CTTAACACAGATTCATAAGCTTGATGAGTTATGATAAAAAGATGTTTCATCATGGAATTTT	226		
Qy	1133	ATGCCCTCTTTT	1145		
Db	225	ATTTCCTTTGTT	213		

```

RESULT 37
5405943-3/c
;Patent No. 5405943
; APPLICANT: COMINGS, DAVID E.
; TITLE OF INVENTION: TOURETTE SYNDROM, AUTISM AND ASSOCIATED
; BEHAVIORS
; NUMBER OF SEQUENCES: 5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/562,596
; FILING DATE: 03-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 125,577
; FILING DATE: 25-NOV-1987
; APPLICATION NUMBER: 271,653
; FILING DATE: 16-NOV-1988
; APPLICATION NUMBER: 410,831
; FILING DATE: 22-SEP-1989
; SEQ ID NO:3
; LENGTH: 1647

```

5405943-3	Query M	
	Best Lo	
	Matches	
	QY	
	Dd	
	QY	
	Dd	
	QY	
	Dd	

```

RESULT 38
US-09-252-991A-8097/c
; Sequence 8097, Application US/09252991A
; Patent NO. 6551795
;
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8097
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8097

```

	Query Match	2.9%	Score 33.8;	DB 4;	Length 1650;
	Best Local Similarity	52.5%;	Pred. No. 2.3;		
	Matches	74;	Conservative	0;	Mismatches 67; Indels 0; Gaps 0;
Qy	358	CCACGGGTAACGGGTCCTGTGCACATTCTTCGATGACATCAGGATCCGCCCGTCT	417		
Db	1456	CCTCACGGAAACGGAAGCCGGTGACCCCGTCGACGATGGTATCGCGAGCGCCCGGTGC	1397		
Qy	418	CATCTGGCGATAACGGGCGCCGGAGACTGACGCTTCAGCCAGTACCATCAAAACGCTT	477		
Db	1396	AGCGCGGATCGCGAGGAGCCGAAGCATGCGCATAGAGCTGACTAGCCCGCAGGGCT	1337		
Qy	478	CATTTTCCGAAGGCATGACCA	498		
Db	1336	CATAACGCGAGGCGCATGAGCA	1316		

RESULT 39
US-09-252-991A-8205
; Sequence 8205, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 8205
; LENGTH: 2145
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8205

Query Match      2.9%; Score 33.8; DB 4; Length 2145;
Best Local Similarity 52.5%; Pred. No. 2.8;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 358 CACGGGTAACAGCGTCCCTGTACATTCTTGAATGACATCAGGATCCCGCGTCT 417
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
717 CCTCAGGAACAGGAAGCGGTGACCCCGTCGACGATGGTATCGGCGAGCGCGCGTGC 776
QY 418 CACTGGCGATAACGGGCACCGCGAGACTGACGCTTCAGCCAGTACCATCAACGCTT 477
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
777 AGCGCGGATCGGCGAGGAGCGAAGCAATTGGCATAGACTGACTGAGCCCGCAGGGCT 836
QY 478 CATTTTCCGAAGCATGACCA 498
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
837 CATAACGCGAGGCGATGAGCA 857
```

```
RESULT 40
US-09-902-540-874
; Sequence 874, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 874
; LENGTH: 8056
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-874
```

```
Query Match      2.9%; Score 33.8; DB 4; Length 8056;
Best Local Similarity 50.3%; Pred. No. 6.7;
Matches 83; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 359 CACGGGTAACAGCGTCCCTGTACATTCTTGAATGACATCAGGATCCCGCGTCTC 418
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 CACGGGACACAGCGGCGCTCTTCTCGTGTGACGATTTCCGACGGGCTCCGGAGCG 192
QY 419 ACTGGCGATACCGGCGACCGCGAGACTGACGTTTCAGCCAGTACCATCAACGCTTC 478
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
193 CGTTGCCACACAGCGGTAGACCGGAGGCCATGGCTCGATGGCGGTGTAGCCAAAGGCTC 252
QY 479 ATTTTCGAGGCGATGACCACTGGCAATCCGGTAGACCGG 523
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
253 GTAGAGCGAAGGACCAAGCGGCCAGATCCGCGATTTCGATGGAGCAG 297
```

Search completed: June 4, 2005, 15:20:32
Job time : 258 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 12:20:41 ; Search time 799 Seconds
(without alignments)
9086.491 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgcagagatggaataaa.....ttttactttttctctgcag 1181

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	64.2	5.4	1134	17 US-10-369-493-46859	Sequence 46859, A
C 2	64.2	5.4	10801	19 US-10-984-449-76	Sequence 76, Appl
C 3	59.2	5.0	2731748	18 US-10-297-465A-1	Sequence 1, Appli
C 4	57.2	4.8	1197	17 US-10-282-122A-26426	Sequence 26426, A
C 5	57.2	4.8	1200	17 US-10-282-122A-28480	Sequence 28480, A
C 6	50.4	4.3	1202	17 US-10-369-493-32569	Sequence 32569, A
C 7	50.2	4.3	1143	17 US-10-282-122A-10729	Sequence 10729, A
C 8	49	4.1	1135	9 US-09-974-300-1695	Sequence 1695, Ap
C 9	48.6	4.1	783	9 US-09-974-300-6118	Sequence 6118, Ap
C 10	48.6	4.1	1179	17 US-10-282-122A-14996	Sequence 14996, A
C 11	48.2	4.1	1113	17 US-10-369-493-44085	Sequence 44085, A

C 12	48.2	4.1	1146	17 US-10-369-493-25021	Sequence 25021, A
C 13	46.6	3.9	1323	15 US-10-156-761-4750	Sequence 4750, Ap
C 14	46.6	3.9	9025608	15 US-10-156-761-1	Sequence 1, Appli
C 15	46.2	3.9	783	9 US-09-974-300-6111	Sequence 6111, Ap
C 16	44.2	3.7	1120	17 US-10-369-493-32590	Sequence 32590, A
C 17	43.4	3.7	1392	15 US-10-156-761-3990	Sequence 3990, Ap
C 18	42.8	3.6	1155	17 US-10-282-122A-27595	Sequence 27595, A
C 19	42.4	3.6	1972	17 US-10-425-114-8	Sequence 8, Appli
C 20	42.4	3.6	1949	17 US-10-425-114-30648	Sequence 30648, A
C 21	42.4	3.6	2077	18 US-10-425-115-47612	Sequence 47612, A
C 22	42.2	3.6	1954	18 US-10-437-963-18583	Sequence 18583, A
C 23	41.6	3.5	1227	9 US-09-974-300-1771	Sequence 1771, Ap
C 24	41.4	3.5	1152	17 US-10-282-122A-13173	Sequence 13173, A
C 25	41.4	3.5	1194	17 US-10-369-493-43673	Sequence 42673, A
C 26	40.8	3.5	363	18 US-10-767-703-17873	Sequence 17873, A
C 27	40.8	3.4	1156	17 US-10-282-122A-9697	Sequence 9697, Ap
C 28	40.6	3.4	836	17 US-10-425-114-1667	Sequence 1667, Ap
C 29	40.6	3.4	2103	18 US-10-425-115-56218	Sequence 56218, A
C 30	40.6	3.4	129908	14 US-10-270-875-1	Sequence 1, Appli
C 31	40.6	3.4	129908	14 US-10-270-878-1	Sequence 1, Appli
C 32	40.6	3.4	129908	14 US-10-270-710-1	Sequence 1, Appli
C 33	40.6	3.4	129908	14 US-10-270-859-1	Sequence 1, Appli
C 34	40.6	3.4	129908	15 US-10-270-845-1	Sequence 1, Appli
C 35	40.4	3.4	1131	17 US-10-282-122A-25777	Sequence 25777, A
C 36	40.4	3.4	2100	17 US-10-369-493-33829	Sequence 33829, A
C 37	39.8	3.4	1209	17 US-10-369-493-24261	Sequence 24261, A
C 38	39.8	3.4	1239	15 US-10-156-761-5343	Sequence 5343, Ap
C 39	39.8	3.4	86114	18 US-10-080-170-648	Sequence 648, App
C 40	39.8	3.4	86114	18 US-10-080-170-648	Sequence 648, App
C 41	39.8	3.4	86114	18 US-10-468-356-648	Sequence 648, App
C 42	39.8	3.4	993	17 US-10-369-493-45361	Sequence 45361, A
C 43	39.6	3.3	1295	18 US-10-425-115-52894	Sequence 52894, A
C 44	39.4	3.3	1244	17 US-10-425-114-34652	Sequence 34652, A
C 45	39.2	3.3			

ALIGNMENTS

RESULT 1

US-10-369-493-46859/c
; Sequence 46859, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46859
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-369-493-46859

Query Match	5.4%;	Score 64.2;	DB 17;	Length 1134;
Best Local Similarity	54.4%;	Pred. No. 1.9e-10;		
Matches 129;	Conservative 0;	Mismatches 108;	Indels 0;	Gaps 0;
Qy	335	GCCTGTCATGCAGACATCCCGGAGTAACAGCGTCCCTGTGCATCTTCTGAAT	394	
Db	990	GGCTGTGCGCGTAAACATCACGACATCCACCAAAATCCGCTCATTTGTTTTAT	931	
Qy	395	GACATCAGGATCCCGCTCTCTACCTGGCGATACGGCAGCCGCGAGACTGACGCTTC	454	

Db 930 AACCTAGGATACCGCAATGTTTGTTCATATCAAGGCACTCCGCAAGCCATCGTTC 871
Qy 455 AGCCAGTACCATACCAACGCTTCATTTTCGGAAGCATGACCAACACACTGGCAATCCG 514
Db 870 AAGCAGGACAAAGCCAAAGCTTCTTTTTCAGATAGCAGAGCTTCAATCGCTAATAGA 811
Qy 515 GTAGACGGTAACGCTGGGAAAGGGACCTGCCATTAAACATCTCCGCTCATTC 571
Db 810 ATAAAGATCTTCAACACAGCGTCTTGATTTCCAAAGCATTAAGACTTGGTCTTCCAGCC 754

RESULT 2

US-10-984-449-76
; Sequence 76, Application US/10984449
; Publication No. US20050089973A1
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Patterson, Thomas A.
; APPLICANT: Hermann, Theron
; TITLE OF INVENTION: METHODS AND MICROORGANISMS FOR PRODUCTION OF
; FILE REFERENCE: BGI-141CPN
; CURRENT APPLICATION NUMBER: US/10/984,449
; CURRENT FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: USSN 09/667,569
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: USSN 09/400,494
; PRIOR FILING DATE: 1999-09-21
; PRIOR APPLICATION NUMBER: USSN 60/210,072
; PRIOR FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: USSN 60/221,836
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: USSN 60/227,860
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 76
; LENGTH: 10801
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
; OTHER INFORMATION: pAN240 plasmid
US-10-984-449-76

Query Match 5.4%; Score 64.2; DB 19; Length 10801;
Best Local Similarity 54.4%; Pred. No. 6.8e-10; Mismatches 108; Indels 0; Gaps 0;
Matches 129; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 335 GCGGTCCATGACAGACATCCCGGGAACAGCGTCCCTGTGTCACATTTCTTGAAT 394
Db 6490 GGCTGTGGCGGTAAACATCACGACATCCACCAAAATCCGCTCACATGTTTTTAT 6549
Qy 395 GACATCAGGATCCCGCGCTCTCATCTGGGATAACGGGACGCGGAGACTGACGGTTC 454
Db 6550 AACCTCAGGATACCGGCAATGTTTGTTCATACAAAGGACCTCCGCAAGCCATCGTTC 6609
Qy 455 AGCCAGTACCATACCAACGCTTCATTTTCCGAAGCATGACCAACACCTGGCAATCCG 514
Db 6610 AAGCAGACAAAGCCAAAGCTTTCTTTTTCAGATAGCAGCAGCTTCAATCGCTAATAGA 6669
Qy 515 GTAGACGGTAACGCTGGGAAAGGGACCTGCCATTAAACATCTCCGCTCATTC 571
Db 6670 ATAAAGATCTTCAACACAGCGTCTTGATTTCCAAAGCATTAAGACTTGGTCTTCCAGCC 6726

RESULT 3

US-10-297-465A-1/c
; Sequence 1, Application US/10297465A
; Publication No. US20040142413A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Andrew

; APPLICANT: Reinach, Fernando
; APPLICANT: Setubal, Joao
; APPLICANT: Medanis, Joao
; APPLICANT: Arruda, Paulo
; TITLE OF INVENTION: Isolated Genome of Xylella Fastidiosa and Uses Thereof
; FILE REFERENCE: FAPESP 202 US (10213376)
; CURRENT APPLICATION NUMBER: US/10/297,465A
; CURRENT FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: PCT/IB01/01618
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/209,906
; PRIOR FILING DATE: 2001-06-17
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2731748
; TYPE: DNA
; ORGANISM: Xylella fastidiosa
US-10-297-465A-1
Query Match 5.0%; Score 59.2; DB 18; Length 2731748;
Best Local Similarity 52.4%; Pred. No. 8.1e-07;
Matches 130; Conservative 0; Mismatches 118; Indels 0; Gaps 0;
Qy 285 GCTGCCATCATCCGGAAGCGTTCCGGTGGGATAAAAAATCGCGCAGTGC CGCGTCCAT 344
Db 837893 GCTGCTGCCATCGCGCCCGGGTTGCGGATACAAAAATGCCAGAAATCGCATTC CGGCAC 837834
Qy 345 GCAGACACATCCCCACAGGGTAAACAGGTCCTCCGTGCATTTCTTGAATGACATCAGGG 404
Db 837833 GCAGGCACGTCCCGCGGGCAGCAGCATCCCGGTCTGTACCATCTCTGCAATGTCTCGCA 837774
Qy 405 ATCCCGCCGCTCTCACTGGCGATAACGGGCACGCCCGAGACTGACGCTTCAGCCAGTACC 464
Db 837773 GCCCACCACATCATCTCACAGCAGGACACACCGGATGCTTGGCGCTCTGGGAGACA 837714
Qy 465 ATACCAACGCTTCATTTTCCGAAGCATGACCAACACACTGGCAATCCGGTAGACCGGT 524
Db 837713 CGTCCAAAGCTCTCCGGCAACAAAGAGCATCGCCAGCATCGCCCAACGCTCGAAATAA 837654
Qy 525 AAGCTGG 532
Db 837653 CGCGCCGG 837646

RESULT 4

US-10-282-122A-26426/c
; Sequence 26426, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA, 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335

```
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26426
; LENGTH: 1197
; TYPE: DNA
; ORGANISM: Mycobacterium bovis
US-10-282-122A-26426

Query Match      4.8%; Score 57.2; DB 17; Length 1197;
Best Local Similarity 54.2%; Pred. No. 5e-08;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 272 TTCTCTGGCGCTGGCTCCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGCGCAG 331
Db 1119 TTCTCGCGCGCGCGCCATCGGACGCGCCCGATCCGATCGATCAGCAACTCGGCCAC 1060

Qy 332 TGGCGCGTCCATGCAGACACATCCCGGTAACAGCGTCCCTGTGCATCTTCTG 391
Db 1059 GGGCTCGGCCACCGCGTCCACCGACTACGTCGACCACTAGCCCACTTTGTGCTG 1000

Qy 392 AATGACATCAGGATCCCGCGTCTCACTGGCGGATAACGGGACGCGGAGACTGACGC 451
Db 999 CACCGTTTCGGCGCTCGCCAGATTCGCGGATTAACGGGACGCGCGGGGAGGC 940

Qy 452 TTCAGCCAGTACCATACCAACGCTTCATTTTC 485
Db 939 TTCGAGGACAGATGCCCAAGCCCTCGAGTCC 906

RESULT 5
US-10-282-122A-28480/c
; Sequence 28480, Application US/10282122A
; Publication No. US2004029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zvekind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
```

```
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 28480
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
US-10-282-122A-28480

Query Match      4.8%; Score 57.2; DB 17; Length 1200;
Best Local Similarity 54.2%; Pred. No. 5e-08;
Matches 116; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 272 TTCTCTGGCGCTGGCTCCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGCGCAG 331
Db 1119 TTCTCGCGCGCGCGCCATCGGACGCGCCCGATCCCGATCGATCAGCAACTCGGCCAC 1060

Qy 332 TGGCGCGTCCATGCAGACACATCCCGGTAACAGCGTCCCTGTGCATCTTCTG 391
Db 1059 GGGCTCGGCCACCGCGTCCACCGACTACGTCGACCACTAGCCCACTTTGTGCTG 1000

Qy 392 AATGACATCAGGATCCCGCGTCTCACTGGCGGATAACGGGACGCGGAGACTGACGC 451
Db 999 CACCGTTTCGGCGCTCGCCAGATTCGCGGATTAACGGGACGCGCGGGGAGGC 940

Qy 452 TTCAGCCAGTACCATACCAACGCTTCATTTTC 485
Db 939 TTCGAGGACAGATGCCCAAGCCCTCGAGTCC 906

RESULT 6
US-10-369-493-32569/c
; Sequence 32569, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 32569
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: Chloroflexus aurantiacus
US-10-369-493-32569

Query Match      4.3%; Score 50.4; DB 17; Length 1202;
Best Local Similarity 51.3%; Pred. No. 1.1e-05;
Matches 117; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Qy 274 CTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGCGCAGTG 333
Db 1104 CTGCGCCACGGGCGCTCGCAGGCGAGACTCGCGCAGGAGAACGACGCTTTGAGCAGGG 1045
```

QY 334 CGCGGTCATGACAGACATCCCGCCACCGGTAAACAGCGTCCCTGTGCACATCTTCTGAA 393
Db 1044 CATGAGCGAGTGTCCGACATCACCTGCGCGCACCAACAGGCCGTTTCAGCTCGTGTGGA 985
QY 394 TGACATCAGGATCCCGCCGTCCTCACTGGGATACGGGACGCGGAGACTGACGCTT 453
Db 984 CAATCTCCGAGGCGCACCAAGTATTCGAGACGATCACTGGGACGCGGCGGCAATGCTT 925
QY 454 CAGCCAGTACATACCAAAAGCTTCAATTTCCGAAGGCATGACCA 501
Db 924 CCAGGGCCACAATGCCAATGGTTCTGACAGCGAAGGGAAGACGGCA 877

RESULT 7

US-10-282-122A-10729/c
; Sequence 10729, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10729
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Bacillus anthracis

US-10-282-122A-10729

Query Match 4.3%; Score 50.2; DB 17; Length 1143;
Best Local Similarity 49.4%; Pred. No. 1.3e-05;
Matches 130; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 241 TTCTGTTAATPATAAACCGGTACTCAATATCTTCTGGCGCTGGCTGCCATCATCCGGA 300
Db 1081 TTTTTCAGAGCGAAACTGCTCGTAAACACTTTCTTCTGCTCTCTCCCATATTAGGT 1022
QY 301 AGCGTTCGGTCCGGATAAAATTCGGCAGTGGCGGTTCATGACACATCCCCCA 360

Db 1021 GAAGTTCTTCATCTCTTTAATAGCTGAATGGCTTGATCTGCCACTCTCTGTTGTATCGCAA 962
QY 361 CGGGTAACAGCGTCCCTGTGCATTTCTCTGAATGACATCAGGGATCCCGCCGTTCTCAC 420
Db 961 CTTTCATATAATATCTGTATCACCATGTTGAATGACTTCTGGAATACCTCCAAACCTTG 902
QY 421 TGGCGATAACGGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCAT 480
Db 901 TCCGATACAGGTACACCACACGCGCATCGCTTCTTAATAAACAAGACCAAAACTCTCCT 842
QY 481 TTTCCGAAGGCATGACCAACCA 503
Db 841 TCTCTGATAAAAGCAACATTAAA 819

RESULT 8

US-09-974-300-1695/c
; Sequence 1695, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1695
; LENGTH: 1135
; TYPE: DNA
; ORGANISM: Bacillus licheniformis

US-09-974-300-1695

Query Match 4.1%; Score 49; DB 9; Length 1135;
Best Local Similarity 50.2%; Pred. No. 3.2e-05;
Matches 121; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 351 ACATCCCCCAGGGTAACAGCGTCCCTGTGCATTTCTCTGAATGACATCAGGGATCCCG 410
Db 977 ATATCTCCGACATCGACCAAGTAACGGATACTCCGTTTTTGTATCACTCGGAAATCCCC 918
QY 411 CCGGTCTCACTGGCGATAACGGGACGCGGAGACTGACGCTTCAGCCAGTACCATACCA 470
Db 917 CCGATATTTGTCCGATGCAACGGAGACGCGGCGGCGCATCGCTTCAAGCAGCACTAAACCG 858
QY 471 AAGCGTTTCATTTTCCGAAGGCATGACCAACACTCGGCAATCCGGTAGACCGGTAAGCT 530
Db 857 AGCTTTCTCTTCGGACAAAAGCAGCTTTAAATCGCTGATGGCGHAAAGCTTTCTACG 798
QY 531 GGGAAAAGGCGACCTGCCATTAAACACATCTCCGCTCAITCCCAAGGTGTTCTGTCTGTA 590
Db 797 CTGTCTTTTTCGAAGGAAAAGAGCTGCTTTTTTAATCCGAGCTGTCTGACAAAGCTG 738
QY 591 C 591
Db 737 C 737

RESULT 9

US-09-974-300-6118/c
; Sequence 6118, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1695
; LENGTH: 1135
; TYPE: DNA
; ORGANISM: Bacillus licheniformis

US-09-974-300-6118

; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974.300
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8491
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6118
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-6118

Query Match 4.1%; Score 48.6; DB 9; Length 783;
Best Local Similarity 49.8%; Pred. No. 3.6e-05;
Matches 123; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 354 TCCCCACGGGTAAACAGCGTCCCTGTCAACATTTCTTGAATGACATCAGGGATCCCGCCC 413
Db 631 TTCCCAAGGGGCAATAATTAACCAAGTTACTTCTGCTTCAACAATTCGGGATGCCACCA 572
QY 414 GTCTCACTGGCGATTAACGGGACGCGCGGAGACTGACGCTTTGAGCCAGTACCATACCAAC 473
Db 571 ATTTTGTGCGATCACAGGTACCCGCGATGCGCTTCAAGGGCCACTAAACCAAG 512
QY 474 GTTTCATTTCCGAGGCGATGACACACACTGGCAATCGGTAGACCGGTAAAGCTGGG 533
Db 511 CTTTCTTTTTCGCTTAACAACAGCATTAATTCGCTCATAGACAACAATTCAGCAACATGC 452
QY 534 AAAAGGCGACTGCGATTAACACATCTCCGCTCATTCAGGCTTCTGCTGCTGACGC 593
Db 451 TTTTGGTGTCTAGCATAGACATCGTTTGTCAAGCCGTTTCTTCCACAAACTGCCTC 392
QY 594 AGACGTG 600
Db 391 GCAATTG 385

RESULT 10
US-10-282-122A-14996/c
; Sequence 14996, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282.122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14996
; LENGTH: 1179
; TYPE: DNA
; ORGANISM: Bordetella pertussis
US-10-282-122A-14996
Query Match 4.1%; Score 48.6; DB 17; Length 1179;
Best Local Similarity 55.7%; Pred. No. 4.5e-05;
Matches 93; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 326 GCGCAGTGGCGGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCTGTCACTT 385
Db 999 GCGCAGCGCCCCCTCGAGGGCAGCCGCTGTCGGAGGAACACGAGCGCCCTCTGCGC 940
QY 386 CTTTCAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAAACGGGACGCGGAGAC 445
Db 939 GTTGGCAACATCTCGTACGCGCGGACATCCAGCGACCCGCGGACCGGACGCGCAGC 880
QY 446 TGACGCTTCAGCGAGTACCATACCAACGCTTCATTTTCGAAAGGCA 492
Db 879 CGCGGCTCGACAAACCGTGCAGCGCTCTTTCGCGGTGGCA 833
RESULT 11
US-10-369-493-44085/c
; Sequence 44085, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369.493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 44085
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Pyrococcus horikoshii
US-10-369-493-44085
Query Match 4.1%; Score 48.2; DB 17; Length 1113;
Best Local Similarity 59.0%; Pred. No. 6e-05;
Matches 102; Conservative 0; Mismatches 68; Indels 3; Gaps 1;
QY 352 CATCCCCACGGGTAAACAGCGTCCCTGTCACTTCTTGAATGACATCAGGATCCCGC 411
Db 976 CATTTCCGGGTGGAACATAAAGCCCTGCACTGTTTCTTTTATAACCTCAGGTATACCAC 917
QY 412 CGCTCTCACTGGCGATAACGGGCAACCGGAGACTGACGCTTCAGCGAGTACCATACCAA 471
Db 916 CAACGTCTGTAGCTATAATGGGNACTCCGGAGCCCATGCCCTCAGGATTACTATGCCGA 857

```
QY 472 ACAGTTCATTTC---CGAAGGCATGACACCACTGGCAATCCGGTAGACC 521
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 856 ACAGTTCGATGATATCGAAGGTAAAGACGACAGCTCGGCATCTCTGAAAACC 804

RESULT 12
US-10-369-493-25021/c
; Sequence 25021, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 39-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25021
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Pyrococcus horikoshii
US-10-369-493-25021

Query Match 4.1%; Score 48.2; DB 17; Length 1146;
Best Local Similarity 59.0%; Pred. No. 6.1e-05;
Matches 102; Conservative 0; Mismatches 68; Indels 3; Gaps 1;

QY 352 CATCCCCCAGGGTAACAGCGTCCCTGTACATCTCTCTCAATGACATCAGGATCCCGC 411
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 985 CATTCGGGTGAACTAAAAAGCCCTGCACTGTTTCTTTATAACCTCAGGTATACCC 926

QY 412 CCGTCTCAGTCGGGATAAACGGGACCGCGAGACTGACGCTTCAGCCAGTACCATACCAA 471
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 925 CAAGCTGTAGTATATATGGAACCTCCGAGGGCCATGGCTCAAGGATTACTATGCCGA 866

QY 472 ACAGTTCATTTC---CGAAGGCATGACACCACTGGCAATCCGGTAGACC 521
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 865 ACAGTTCGATGATATCGAAGGTAAAGACGACAGCTCGGCATCTCTGAAAACC 813

RESULT 13
US-10-156-761-4750/c
; Sequence 4750, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4750
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (1)...(1323)
US-10-156-761-4750

Query Match 3.9%; Score 46.6; DB 15; Length 1323;
Best Local Similarity 47.9%; Pred. No. 0.00024;
Matches 167; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 274 CTCTGGCGCTGGTGGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCCGCAGTG 333
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1099 CGGGGCGCGGGCCCATCGACCTCGGCAGATCGTCTCCAGGACACAGGTCAGGATCG 1040

QY 334 CGCCGGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCTGTGACATTTCTTGAA 393
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1039 CGTCGGCGAGGGCGGGGTGTGCGGCACCGGCACAGCGCGCGTCCACGCGTCCGTGA 980

QY 394 TGACATCAGGATCCCGCCCG---TCTCACTGGCGATACGGGCAGCCGGAGACTGACG 450
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 979 TGATCTCGCGGGGCGGAGGGGGCAGTCCGTGCTGACCAACCCCATCAGCGGGCTGACCGGCATCG 920

QY 451 CTTACGCCAGTACATACCAAAACGCTTCATTTTCCGAAGGCATGACACCACTGGCAA 510
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 919 CCTGACACAGCGTCATCCCGAACGACTCCGCGTCCGAGGGCTGACACACAGGACGCT 860

QY 511 TCCGCTAGACCGGTAAACGCTGGGAAAAGGCACCTGCCATTAAACATCTCCGCTCATTC 570
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 859 GGGCGAACTCCGCTCGATGGCGGTACGCACCCCATCAGCGGGCTGACCGCGGAGCC 800

QY 571 CGAGGTGTTCTGTCTGTGACGACGAGCGTGTCTGTTATTTCTTACAGGCC 619
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 799 CCAACCCGTCGATCAGCCCTCGAACCCGCTCTTCTCGGCGCGCGCTCC 751

RESULT 14
US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 3.9%; Score 46.6; DB 15; Length 9025608;
Best Local Similarity 47.9%; Pred. No. 0.035;
Matches 167; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 274 CTCTGGCGCTGGTGGCCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCCGCAGTG 333
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5809166 CGGGGCGCGGGCCCATCGACCTCGGCAGATCGTCTCCAGGACACAGGTCGAGCATCG 5809225

QY 334 CGCCGGTCCATGCAGACACATCCCCACGGGTAAACAGCGTCCCTGTGACATTTCTTGAA 393
||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

Db 5809226 CGTCCGGAGGGCCGGGTTCGCGCAGGGGACACAGCCGGCCCTCCACGGGTCCTGTA 5809285
Qy 394 TGACATCAGGATCCCGCCCG--TCTCACTGGCGATAACGGGACACGGGAGACTGACG 450
Db 5809286 TGATCTCGGGGGCCGAGGGGAGTCTGCTGACACGGGACACCGCCGAGGCATCG 5809345
Qy 451 CTTTCAGGAGTACATACCAACGCTTCAATTTTCGGAAGGATGACACCACTGGCAA 510
Db 5809346 CCTCGACAGCGTTCATCCGAACGACTCCGGCTCCGAGGCGCTGACCAACGAGCGCT 5809405
Qy 511 TCCGGTAGACGGTAACGCTGGGAAAGGACCTGCGCATTAACACATCTCCGCTCATTC 570
Db 5809406 GGGCGAATCTCGCTCGATGGGCTAGCACCCTCCATAGCCGGGCTGACCGCGAGCC 5809465
Qy 571 CCAGGTCTTCTGCTGTCGACGACGATGCTTCTGTAATTTTCACGCCC 619
Db 5809466 CCAACCGTCGATCAGCCCTCGCAACGCTCTTCTCGGCCCGCCCTCC 5809514

```

```

RESULT 15
US-09-974-300-6111/c
; Sequence 6111, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Betka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6111
; LENGTH: 783
; TYPE: DNA
; ORGANISM: Bacillus clausii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(783)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-6111

```

```

Query Match 3.9%; Score 46.2; DB 9; Length 783;
Best Local Similarity 59.5%; Pred. No. 0.00024;
Matches 78; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 354 TCCCCACGGGTAAACAGCGTCCCTGTGACATTTCTTGAATGACATCAGGGATCCCGCCC 413
Db 633 TTCCCAAGGGGCAATAAATACCGTTACTTCTTCAACAATTCGCGGATGCCACCA 574
Qy 414 GTTCTACTGGCGATAACGGGACGCCGAGACTGACGCTTTCAGCCAGTACCAATCAAC 473
Db 573 ATTTTGTGCGATCACAGGTACCCGCGATGCGCTTCAAGCGGCACTAAACCAAG 514
Qy 474 GCTTCATTTTC 484
Db 513 CTTTCTTTTTC 503

```

```

RESULT 16
US-10-369-493-32590/c
; Sequence 32590, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

```

```

; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 32590
; LENGTH: 1120
; TYPE: DNA
; ORGANISM: Chloroflexus aurantiacus
; US-10-369-493-32590

```

```

Query Match 3.7%; Score 44.2; DB 17; Length 1120;
Best Local Similarity 50.2%; Pred. No. 0.0014;
Matches 109; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 285 GGTGCCATCATCCGGAAGCGTTCCGTCGGGATAAAAAATCGGCGAGTGGCCGGTCCAT 344
Db 1035 GCGGCCAGCGCTGCCCGACGTGCCGATCGAGAGCAGACTATTGACAGCCGCGAGCCAGC 976
Qy 345 GCAGACACATCCCCACGCGGTAAACAGGTCCTGTGACATTTCTTGAATGACATCAGG 404
Db 975 GCGGCCGATCGCGAGGCGGTACCACTAGTACCGGTTCTGACCCCTCGGCGAGACTTCGGGA 916
Qy 405 ATCCCGCCGCTCTCACTGCGGATAACGGGACACCGGAGACTGACGCTTCAGCCAGTACC 464
Db 915 AAGCCTCCGAAAGGGTTGCTACAAACCGGTAATCCGACGCGCTTGCACGCGGCCA 856
Qy 465 ATACCAACGCTTCATTTTCCGAAAGCATGACCA 501
Db 855 ATGCCAAAGCTTTCGCTGCGCATAGCTGGTTGCCAGCA 819

```

```

RESULT 17
US-10-156-761-3990/c
; Sequence 3990, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 3990
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1392)
US-10-156-761-3990

```

```

Query Match 3.7%; Score 43.4; DB 15; Length 1392;
Best Local Similarity 49.3%; Pred. No. 0.0031;
Matches 113; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

Qy 289 CCATCATCCGGAAGCGTTCCGTCGGGATAAAAAATCGGCGAGTGGCCGGTCCATGCGAG 348

```

Db 1273 CCATCCGGCGCAGACAGCGCGGGTTCGTCGCGAAATCGCGAGTACGCGCGCGTAGGCGG 1214
QY 349 ACACATCCCGCCAGGGTAACAGGTCCTCTGCACATTTCTTCTGAATGACATCAGGGATCC 408
Db 1213 CGGGATCGTCCCTGTGACGAGAAATCCGTCCGCGGTGCGGACCGCGACCGGAAGCC 1154
QY 409 CGCCCGCTCTCACTGGCGGATACCGGCAACGCGGAGACTGACGGTTTCAGCCAGTACCATAC 468
Db 1153 CGCCACGAGGCGCGCAGCACCGCGTACCGGCGCGCTCTATGGCGACGACGC 1094
QY 469 CAAACGCTTCATTTTCGAAAGGATGACACACACTGGCAATCCGGTA 517
Db 1093 CGAACGACTCGTTGTAGGAGGCATGACACGAGCGGACGCGCGCGGAA 1045

RESULT 18

US-10-282-122A-27595/c
; Sequence 27595, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27595
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Mycobacterium leprae
US-10-282-122A-27595

Query Match 3.6%; Score 42.8; DB 17; Length 1155;
Best Local Similarity 50.5%; Pred. No. 0.0044;
Matches 104; Conservative 0; Mismatches 102; Indels 0; Gaps 0;
QY 280 CGCTGGCTGCATCATCCGGAAGGTTCCGGTCGGGATAAAAAATCGGCAGTGCAGCGG 339
Db 1069 CGGCTGGCCCATAGCAGCGCCCGGTACAGTCGGTCAGTAATCTCGTGACTGCGTCAG 1010

QY 340 TCCATGCGAGACATCCCGCCAGGGTAACAGCGTCCCTGTGCACATTTCTTCTGAATGACAT 399
Db 1009 CCACCATATTTCAACCGACCGCGCTCAACTACTAGTCCAGTCTTGTGTCGACAGCTT 950
QY 400 CAGGGATCCCGCGCTCTCACTGGCGGATAAACGGGCAACCGCGAGACTGAGCGTTCAAGCCA 459
Db 949 CGGAGCTCCGCGCGGATTTTCCAGCGATTAACCGGCAACCGCGGTAGCCGAGGCTCCAGGA 890
QY 460 GTACCATACCAACGCTTCATTTTCC 485
Db 889 ACAGATGCCCAACCTCGAGCTCC 864

RESULT 19

US-10-425-114-8/c
; Sequence 8, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 8
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700041786_FLI
US-10-425-114-8

Query Match 3.6%; Score 42.4; DB 17; Length 972;
Best Local Similarity 54.5%; Pred. No. 0.0055;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 371 CGTCCCTGTCACATTTCTTGAATGACATCAGGATCCCGCGTCTCACTGGCGATAAC 430
Db 295 CCTCCCTCTGTATGATATCAGTATTCTCCGCGCGAGCCCGACTAC 236
QY 431 GGGCAGCGCGGAGACTGACGCTTCAGCCAGTACCACCAAGCTTCAATTTTCCGAAG 490
Db 235 CGGAATCTCTGATGATCGCTCCAGCACCAACCCCACTGTCTCTGACTCGGAAG 176
QY 491 CATGACCAACACTGGCAATCCGCTAGACCGGTAA 526
Db 175 CATCAACAACATATCCCACTGGCATAGGCTGTGA 140

RESULT 20

US-10-425-114-30648/c
; Sequence 30648, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 30648

```
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLB73055B03_FLI
US-10-425-114-30648

Query Match      3.6%; Score 42.4; DB 17; Length 1949;
Best Local Similarity 54.5%; Pred. No. 0.0082;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGATCCCGCCGTCTCTCACTGGCGATAAC 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1269 CTTCCCTCTGTATCTTCAGTATGATATCAGGTATTCCTCCGGGGAGCCCGGACTAC 1210
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 431 GGGCAGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAAAGCTTCATTTTCCGAAGG 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1209 CGGAACCTCCTGATGACATGGCTCCAGCACCACAAACCCAGTGCTCTGACTCGGAAGG 1150
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 491 CATGACACACACTGGCAATCCGGTAGACCGGTAA 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1149 CATCAAAACACATCCCCCACTGGCATAGGCTGTGA 1114
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 21
US-10-425-115-47612/c
; Sequence 47612, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 47612
; LENGTH: 2077
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_14342C.1
US-10-425-115-47612

Query Match      3.6%; Score 42.4; DB 18; Length 2077;
Best Local Similarity 54.5%; Pred. No. 0.0085;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGATCCCGCCGTCTCTCACTGGCGATAAC 430
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1328 CTTCCCTCTGTATCTTCAGTATGATATCAGGTATTCCTCCGGGGAGCCCGGACTAC 1269
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 431 GGGCAGCCGGAGACTGACGCTTCAGCCAGTACCATACCAAAAGCTTCATTTTCCGAAGG 490
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1268 CGGAACCTCCTGATGACATGGCTCCAGCACCACAAACCCAGTGCTCTGACTCGGAAGG 1209
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 491 CATGACACACACTGGCAATCCGGTAGACCGGTAA 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1208 CATCAAAACACATCCCCCACTGGCATAGGCTGTGA 1173
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 22
US-10-437-963-18583
; Sequence 18583, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
```

```
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 18583
; LENGTH: 1954
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24127C.1
US-10-437-963-18583

Query Match      3.6%; Score 42.2; DB 18; Length 1954;
Best Local Similarity 55.9%; Pred. No. 0.0096;
Matches 80; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 384 TTCTTCTGAATGACATCAGGATCCCGCCGTCTCTCACTGGCGATAACGGGACGCCCGGAG 443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 579 TCCTTGGGTATTATATCAGGTATACCTCCAGCAGCAGCAGCAGCAACTGGGACTCCAGAA 638
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 444 ACTGAGCTTCAGCCAGTACCATACCAAAAGCTTCATTTTCCGAGGATGACCACCACA 503
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 GCCATGGACTCCAGCACTACTTGACCAAGTGTCTCAGATTCTGAAGGCATTGCAAAACAAG 698
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 504 CTGGCAATCCGGTAGACCGGTAA 526
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 699 TCGCCACTGGCATATGCTGTGA 721
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 23
US-09-974-300-1771/c
; Sequence 1771, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10095.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1771
; LENGTH: 1227
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1771

Query Match      3.5%; Score 41.6; DB 9; Length 1227;
Best Local Similarity 62.5%; Pred. No. 0.012;
Matches 65; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 352 CATCCCCCAGGGTAAACAGCGTCCCTGTCACTTCTTCAATGACATCAGGGATCCCGC 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1072 CATCTCTTTAGCGGTACCACATACCGGTCACTTATGTTCAACGATTTCGGGAAGTCCCC 1013
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 412 CGGTCTCACTGGCGATAACGGGACGCCCGGAGTACGCTTCA 455
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1012 CTGCATCGCTGACGATAATAGGAAGTCCCGCTAATTGCGCTTCA 969
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 24
```

US-10-282-122A-13173/c

; Sequence 13173, Application US/10282122A
; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.

APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 13173

; LENGTH: 1152

; TYPE: DNA

; ORGANISM: Burkholderia fungorum

US-10-282-122A-13173

Query Match 3.5%; Score 41.4; DB 17; Length 1152;

Best Local Similarity 46.2%; Pred. No. 0.013;

Matches 138; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 279 GCGCTGCTGCATCATCGGAAGCGTTCCGTCGGGATATAAAATCGCGCATGCGCGG 338

Db 980 GCGCGCGGTCAGCGCGCATCGCGGTCATGAGATCGCGCATCGCTTC 921

QY 339 GTCATGACACATCCCCACGGGTAAACAGCGTCCCTGTCACTTCTTGAATGACA 398

Db 920 GCAATGACGGGATCTTGACCGGACAGTACCGCGCTTCGCGCGGTCGACATC 861

QY 399 TCAGGGATCCGCGCTCTCACTGGGGATAACGGGCACGCGGAGACTGACGCTTCAGCC 458

Db 860 TCGGGAATCGCGCCACGCGCTCGCACAGTCGCGGCCAGCTCACGCGCTTCGCAC 801

QY 459 AGTACATACCAACCGCTTCATTTTCGAAGGATGACACACACTGGCAATCCGGTAG 518

Db 800 AGCACAGCCGCGCGCTTCATAGTGAAGGAGCAGCACACATCGCGCTGGGGGAGA 741

QY 519 ACCGGTAAGCTGGGAAAGGGACCTGCCATTAACACATCTCCGCTCATTCGCGGTG 577

Db 740 TAGCGACGGGTCCCGGACAAACCTTGAAACGCACTTCGCGCGGACACCCAGCTG 682

RESULT 25

US-10-369-493-42673/c

; Sequence 42673, Application US/10369493

; Publication No. US20030233675A1

GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 42673

; LENGTH: 1194

; TYPE: DNA

; ORGANISM: Anabaena PCC7120

US-10-369-493-42673

Query Match

Best Local Similarity 48.9%; Pred. No. 0.014;

Matches 111; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 254 AAACCGGTACTCAATATCTTCTGCGCTGGGTCCATATCCGGAAGCTTCGGTGC 313

Db 1119 AAAGCGATGTTCTAAATCTTATAGGCGTTATCAATCAGCCATTGACTATAACCCGATT 1060

QY 314 GGTAAAAATCGCGAGTGGCGGTCCATCGACACATCCCCACGGTTAACGCT 373

Db 1059 TTTCAAACTTCTAGAAATCCCAAGCCAAAGATCGTGGTTATTCACCCAAAGTTACGAT 1000

QY 374 CGCTGTCAATTTCTGTAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAACGGG 433

Db 999 ACCTGTTCGGGTATGTTGCATCTACCTCAGGGAAGCGCTGTATCGGATACCACTACGGG 940

QY 434 CAGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTCAT 480

Db 939 AACCCGCGAGGCGAAGCTTTCTAAAGCTACAATCCCAAGGGTTTCAT 893

RESULT 26

US-10-767-701-17873/c

; Sequence 17873, Application US/10767701

; Publication No. US20040172684A1

GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B

; CURRENT APPLICATION NUMBER: US/10/767,701

; CURRENT FILING DATE: 2004-01-29

; NUMBER OF SEQ ID NOS: 63128

; SEQ ID NO 17873

; LENGTH: 363

; TYPE: DNA

; ORGANISM: Sorghum bicolor

FEATURE:

; OTHER INFORMATION: Clone ID: LIB3476-052-P1-K1-C9

US-10-767-701-17873

Query Match

Best Local Similarity 55.7%; Score 40.8; DB 18; Length 363;

Matches 78; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 387 TTCTGAATGACATCAGGGATCCCGCGTCTCACTGGCGATAACGGGACCGCGGAGACT 446

Db 327 TTGGGTATTATATCAGGTATCCCTCCAGCAGCAGCAACTGGGACTCCAGAGCC 268
Qy 447 GACGCTTCAGCAGTACCATACCAACGCTTCATTTTCCGAAGCATGACCAACCACTG 506
Db 267 ATGACTCCAGCACTACTTGGCCGAGGGTCTCAGACTCTGAAGCATTCGAATACATCT 208
Qy 507 GCATCCGGTAGACCGGTAA 526
Db 207 GCATGGCGTAAGCTTGTGA 188

RESULT 27

US-10-282-122A-9697/c
; Sequence 9697, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A US/10/282,122A
; CURRENT APPLICATION NUMBER: 60/191,078
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-09-06
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9697
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Bacillus anthracis
US-10-282-122A-9697

Query Match 3.5%; Score 40.8; DB 17; Length 1156;
Best Local Similarity 49.6%; Pred. No. 0.022;
Matches 131; Conservative 0; Mismatches 132; Indels 1; Gaps 1;
Qy 241 TTCTGTTAATCAACCGGTACTCAATATCTTCTTGGCGCTGGCTGCCATCATCCGGA 300
Db 1103 TTTTTCAGAGCGAACTGCTCGTAACACTTTCTTGTCTGCTCTCCCATATTACGGT 1044
Qy 301 AGCGTTCGGTCGGGATAAAATCGCGAG-TCCGCCGCTCCATGACACATCCCC 359
Db 1043 GAAGTTCCTACATCTTTTAATAGCTGAATGGCTTGATCTGCCACTCCTGTGTATGCCA 984

Qy 360 ACGGTAACAGCGTCCCTGTTCACATTTCTTGAATGACATCAGGATCCCGCGCTCTCA 419
Db 983 ACTTTCATATAAATATCTGTATCACCATGTTGAATGACTTCTGGAATACCTCCAAACCTT 924
Qy 420 CTGGCATACAGCGGACGCGGAGACTGACGCTTTCAGCCAGTAGTACCATACCAACGCTTCA 479
Db 923 GTTCCGATACAGGTACACACAGGCATCGCTTCTTAATAAACAAGACCAAACTCTCC 864
Qy 480 TTTTCCGAAGGCATGACCAACCA 503
Db 863 TTCTCTGATAAAAGCAACATTAAA 840

RESULT 28

US-10-425-114-1667/c
; Sequence 1667, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1667
; LENGTH: 836
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700163822_FLI
US-10-425-114-1667

Query Match 3.4%; Score 40.6; DB 17; Length 836;
Best Local Similarity 55.2%; Pred. No. 0.021;
Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

Qy 384 TTCTTCTGAATGACATCAGGATCCCGCGCTCTCTACTGGCGATAACGGGACGCCGGAG 443
Db 385 TCCTTGGGTATTATATCAGGTATCCCTCCAGCAGCAGCAGCAGCTGGGACTCCAGAA 326
Qy 444 ACTGACGCTTCAGCCAGTACCATACCAACGCTTCATTTTCCGAAGGCATGACCAACCA 503
Db 325 GCCATGGACTCCAGCACTACTTCCCGAGGGTCTCGGATTTCTGAAGCATTTGCAATACA 266
Qy 504 CTGGCAATCCGGTAGACCGGTAA 526
Db 265 TCCGCACTGGCGTACGCTTGTGA 243

RESULT 29

US-10-425-115-56218/c
; Sequence 56218, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 56218
; LENGTH: 2103

```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_151262C.1
US-10-425-115-56218

Query Match          3.4%; Score 40.6; DB 18; Length 2103;
Best Local Similarity 55.2%; Pred. No. 0.036;
Matches 79; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 384 TTCTTCTGAATGACATCAGGAGTCCCGCGTCTCACTGGCGATAACGGGACGCCGGAG 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 TCCTTGGTATTATATCAGTATCCCTCCAGCAGGACGAGACTGGGATCCAGAA 1431

QY 444 ACTGACGCTTCAGCAGTACCATACCAACGCTTCATTTTCCGAAGCATGACCAACACA 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1430 GCCATGGACTCCAGCACTACTTGGCGAGGGTCTCGGATTCTGAAGGCAATGCAATACA 1371

QY 504 CTGGCAATCCGGTAGACCGGTAA 526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1370 TCCGCACCTGGCGTACGCTTGTA 1348

RESULT 30
US-10-270-875-1/c
; Sequence 1, Application US/10270875
; Publication No. US20030082741A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,875
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-875-1

Query Match          3.4%; Score 40.6; DB 14; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.37;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGTCTCACTGGGATAACGGGACGCCGGAGACTGACGCTTCAGC 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10181 ATCTGCAATACCAACCGCTCATAGTAGCAACACCGGGGTCCACACAGCGACGCTCAAG 10122

QY 458 CAGTACCATACCAACGCTTCATTTCCGAAGCATGACCAACCACTGGC 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10121 CGAAGCGATCCCGAACCCCTCATAGATGAAGCTGCAGCACCACATCGGC 10071

RESULT 31
US-10-270-878-1/c
; Sequence 1, Application US/10270878
; Publication No. US20030082790A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
```

```
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,878
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-878-1

Query Match          3.4%; Score 40.6; DB 14; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.37;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGTCTCACTGGGATAACGGGACGCCGGAGACTGACGCTTCAGC 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10181 ATCTGCAATACCAACCGCTCATAGTAGCAACACCGGGGTCCACACAGCGACGCTCAAG 10122

QY 458 CAGTACCATACCAACGCTTCATTTCCGAAGCATGACCAACCACTGGC 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10121 CGAAGCGATCCCGAACCCCTCATAGATGAAGCTGCAGCACCACATCGGC 10071

RESULT 32
US-10-270-786-1/c
; Sequence 1, Application US/10270786
; Publication No. US20030087392A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjorleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjonsson
; APPLICANT: Arnthor Aevarsson
; APPLICANT: Jakob K. Kristjansson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,786
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-786-1

Query Match          3.4%; Score 40.6; DB 14; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.37;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 398 ATCAGGATCCCGCGTCTCACTGGGATAACGGGACGCCGGAGACTGACGCTTCAGC 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10181 ATCTGCAATACCAACCGCTCATAGTAGCAACACCGGGGTCCACACAGCGACGCTCAAG 10122

QY 458 CAGTACCATACCAACGCTTCATTTCCGAAGCATGACCAACCACTGGC 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 10121 CGAAGCGATCCCGAACCCCTCATAGATGAAGCTGCAGCACCACATCGGC 10071

RESULT 33
US-10-270-710-1/c
; Sequence 1, Application US/10270710
; Publication No. US20030092128A1
; GENERAL INFORMATION:
```



```

; APPLICANT: Sigrður Hjørleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,710
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-710-1

Query Match          3.4%; Score 40.6; DB 14; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.37;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 398 ATCAGGATCCCGCGCTCTCACTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGC 457
Db 10181 ATCTGCAATACCAACCGCTCATAGTACCAACCAACCGGGCTCCACACAGCGCGCTCAAG 10122

Qy 458 CAGTACCATACCAACCGCTTCATTTCCGAGGATGACACCACTGGC 508
Db 10121 CGAAGCGATCCCGAACCCCTCATTTAGATGAAGCTGCAGCACCATCGGC 10071

RESULT 34
US-10-270-859-1/c
; Sequence 1, Application US/10270859
; Publication No. US20030092134A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjørleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,859
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-859-1

Query Match          3.4%; Score 40.6; DB 14; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.37;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 398 ATCAGGATCCCGCGCTCTCACTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGC 457
Db 10181 ATCTGCAATACCAACCGCTCATAGTACCAACCAACCGGGCTCCACACAGCGCGCTCAAG 10122

Qy 458 CAGTACCATACCAACCGCTTCATTTCCGAGGATGACACCACTGGC 508
Db 10121 CGAAGCGATCCCGAACCCCTCATTTAGATGAAGCTGCAGCACCATCGGC 10071

RESULT 35
US-10-270-846-1/c
; Sequence 1, Application US/10270846
; Publication No. US20030129727A1
; GENERAL INFORMATION:
; APPLICANT: Sigrður Hjørleifsdóttir
; APPLICANT: Gudmundur O. Hreggvidsson
; APPLICANT: Olafur H. Fridjónsson
; APPLICANT: Arnthor Aevársson
; APPLICANT: Jakob K. Kristjánsson
; TITLE OF INVENTION: Bacteriophage RM378 of a Thermophilic
; TITLE OF INVENTION: Host Organism
; FILE REFERENCE: 2739.1001-001
; CURRENT APPLICATION NUMBER: US/10/270,846
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/585,858
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 60/137,120
; PRIOR FILING DATE: 1999-06-02
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 129908
; TYPE: DNA
; ORGANISM: Bacteriophage RM378
US-10-270-846-1

Query Match          3.4%; Score 40.6; DB 15; Length 129908;
Best Local Similarity 60.4%; Pred. No. 0.37;
Matches 67; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 398 ATCAGGATCCCGCGCTCTCACTGGCGATAACGGGCACGCCGGAGACTGACGCTTCAGC 457
Db 10181 ATCTGCAATACCAACCGCTCATAGTACCAACCAACCGGGCTCCACACAGCGCGCTCAAG 10122

Qy 458 CAGTACCATACCAACCGCTTCATTTCCGAGGATGACACCACTGGC 508
Db 10121 CGAAGCGATCCCGAACCCCTCATTTAGATGAAGCTGCAGCACCATCGGC 10071

RESULT 36
US-10-282-122A-25777/c
; Sequence 25777, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
```

```
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25777
; LENGTH: 1131
; TYPE: DNA
; ORGANISM: Mycobacterium avium
; US-10-282-122A-25777

Query Match      3.4%; Score 40.4; DB 17; Length 1131;
Best Local Similarity 49.5%; Pred. No. 0.029;
Matches 104; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 276 CTGCGCTGCTGCCATCATCCGAAGCGTTCCGGTCCGGATAAAAATCCGCGAGTGC 335
Db 1064 CGCGCGCGCGCCCATGCGCGCGCGCGTCCGGTCCGTGAGCAATTCGGTGACGGCG 1005

QY 336 CCGGTCAATCAGACATATCCCAACGGGTAAACAGGTCCCTGTCAATCTTCTGAATG 395
Db 1004 TCGGGATCTCGGTACCGAGCGCGCTCCACCACAGCCCGGTCTTTGTGTGTGAC 945

QY 396 ACATCAGGATCCGCGCGCTCTCACTGGCGATAACGGGCACGCGGAGACTGACGTTCA 455
Db 944 GTTTCGGTCTCCCGGACCTTCCCGGCAATCACCGGACACGCTGGCGGAGGCTCC 885

QY 456 GCCAGTACCATAACAAACGCTTCATTTTC 485
Db 884 AGGTAGACGATGCCAGCCCTCGACGTCC 855

RESULT 37
US-10-369-493-33829/c
; Sequence 33829, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33829
; LENGTH: 2100
; TYPE: DNA
; ORGANISM: magnetite-containing magnetic coccus
; US-10-369-493-33829

Query Match      3.4%; Score 40; DB 17; Length 2100;
Best Local Similarity 51.1%; Pred. No. 0.057;
Matches 94; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 321 AATCGCGAGTGGCGGTCCTGATGAGACATCCCGGTAACACGCTCCCTGTC 380
Db 1205 ATATCATCAATGCTCGCAATGGCTCCCGCAACCGGCTATGACGACGGCATTT 1146

QY 381 ACATTTCTTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAACGGGCGCGC 440
```

```
Db 1145 TTGCAATTGAAACAATATCAATCGCGCCCATCTCGGTGGCCACAGGGGCAATCCG 1086
QY 441 GAGACTGAGCTTCCGCCAGTACCATACCAAGCGTTTCATTTTCCGAAGCATGACCA 500
Db 1085 CAAGCCGCGCTTCGATTAGCGTTAGCCCAAAAGGCTCGATCAAGGCGGGTTAACAA 1026
QY 501 ACAC 504
Db 1025 ACAC 1022

RESULT 38
US-10-369-493-24261/c
; Sequence 24261, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 24261
; LENGTH: 1209
; TYPE: DNA
; ORGANISM: Deinococcus radiodurans
; US-10-369-493-24261

Query Match      3.4%; Score 39.8; DB 17; Length 1209;
Best Local Similarity 51.4%; Pred. No. 0.049;
Matches 92; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 344 TCGAGACATCCCCACCGGTTAACAGCGTCCCTGTACATTTCTTGATGATCAGG 403
Db 969 TTGGGACAGTCCGCGCGCGCGCTCAGATACCGGCTGTAGCCGCTTTGCACGCGGTCAG 910
QY 404 GATCCCGCGCTCTCACTGCGGATAACGGGCACGCGGAGACTGAGCTTTCAGCCAGTAC 463
Db 909 GGTGCGCGCGCGCGCCACCGGACCGGACCGCCCATCACTGCGCTTTTTCGACGAC 850
QY 464 CATACCAACGCTTTCATTTTCCGAAGGCGATGACACACACATCGGCAATCCGGTAGACCG 522
Db 849 TAAGCCTGGGTTTCGGTGTCCGAGGCGAAGACGACGAGCGGTAGTAGTACG 791

RESULT 39
US-10-156-761-5343/c
; Sequence 5343, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 11:47:01 ; Search time 4423 Seconds
(without alignments)
10163.668 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgcagagatggaataaa.....ttttactttttctgtgcag 1181

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55.4	4.7	1150	8	BZ567774
2	54.6	4.6	945	8	BZ569481
3	54	4.6	903	8	BZ569276
4	53	4.5	920	8	BZ569436
5	49.2	4.2	890	8	BZ569367
6	48.8	4.1	755	8	AQ989708
7	48.2	4.1	608	6	CD933246
8	48.2	4.1	648	6	CD909568
9	48.2	4.1	887	7	CK153896
10	46.8	4.0	1872	9	CL982059
11	46.6	3.9	328	6	CA735693
12	46.6	3.9	507	6	CD881543
13	46.6	3.9	569	4	BJ544067
14	46.6	3.9	609	4	BJ551686
15	46.6	3.9	615	1	AV909088
16	46.6	3.9	629	4	BJ264853
17	46.6	3.9	640	4	BJ546361
18	46.6	3.9	641	4	BJ545036
19	46.6	3.9	647	4	BJ552866
20	46.6	3.9	651	4	BJ546519
21	46.6	3.9	669	1	AV910971
22	46.6	3.9	683	4	BJ551738
23	46.6	3.9	693	1	AV945422
24	46.6	3.9	713	4	BJ552083

25	45.8	3.9	355	9	AG214640	AG214640	ORYZA sat
26	45.8	3.9	647	4	BJ312175	BJ312175	BJ312175
27	45.8	3.9	725	6	CB663932	CB663932	OSJNE09G
28	45.8	3.9	736	6	CB665280	CB665280	OSJNEG11J
29	45.6	3.9	538	6	CB212014	CB212014	OML02294
30	45.4	3.8	313	7	CF587963	CF587963	USDA-FP_1
31	45.4	3.8	313	7	CN583481	CN583481	USDA-FP_1
32	45	3.8	495	5	BQ763870	BQ763870	EBR003_SQ
33	44.4	3.8	566	4	BJ544878	BJ544878	BU544878
34	43.4	3.7	887	7	CK153409	CK153409	FGAS03201
35	42.4	3.6	281	9	CG058902	CG058902	PUBBY81TD
36	42.4	3.6	431	6	CA300278	CA300278	SCRFLV103
37	42.4	3.6	528	6	CA107818	CA107818	SCRHR107
38	42.4	3.6	599	6	CA169385	CA169385	SCJFSB101
39	42.4	3.6	616	6	CA080408	CA080408	SCVPAM105
40	42.4	3.6	631	6	CA188291	CA188291	SCCCAM100
41	42.4	3.6	659	6	CA169470	CA169470	SCJFSB101
42	42.4	3.6	732	6	CA213271	CA213271	SCJGSB113
43	42.4	3.6	740	6	CA075448	CA075448	SCJLAW106
44	42.4	3.6	769	6	CA106257	CA106257	SCJLHR102
45	42.4	3.6	775	9	CL174179	CL174179	104_378_1

ALIGNMENTS

RESULT 1
BZ567774
LOCUS pac82-164_7205.y2 pac82-164 pseudomonas aeruginosa genomic clone
DEFINITION pac82-164_7205.y2 pac82-164 pseudomonas aeruginosa genomic clone
ACCESSION BZ567774
VERSION BZ567774.1 GI:27199602
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1150)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES

Location/Qualifiers
1..1150
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pac82-164_7205"
/clone_lib="pac82-164"
/note="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match 4.7%; Score 55.4; DB 8; Length 1150;
Best Local Similarity 49.5%; Pred. No. 1.1e-05;
Matches 143; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
QY 234 TGTGCGGTTCTGTATATCAACCGGTACTCAATATCTTCTGTGGCGTGGTGCATC 293
|||
Db 374 TGGACGATATCTTCGATCGCAAAATCAGCTCTGCGCAACGCTGCCGAGCGTCCCATC 433
|||
QY 294 ATCCGAAGCGTTCGCGGTGGGTAATAAATCGCGCAGTCCGCGGTCCATGACAGACA 353

Db 434 TTCTTGGCGCAACTCTGGCGGACTCAATAGACGCTGTAATGGCATCCGCGACAGCAGAGAA 493
QY 354 TCCCCACAGGTAACAGCGTCCCTGTGCACATTCTTCTGAATGATCATCAGGATCCCGCCC 413
Db 494 TCAGCACCAGCAGCAGCGCCGTAACATCAGCTGGATTCATCCCGCAACCCGGT 553
QY 414 GTCTCACTGGCGATAACGGGCAGCGCCGAGACTGACGCTTCAGCCAGTACCATACCAAAAC 473
Db 554 ACATGGTGTCTCAACACCGCAGCCGCGACGCTCGACAGGACTTGGGAGC 613
QY 474 GCTTCACTTCCGAGGCGATGACACCACTGGCAATCCGGTAGACCG 522
Db 614 CTTTCAGTAAGAAGGAGCAGACGACTATATGCGAACGGCGGAAGACAG 662

RESULT 2
BZ569481/c
LOCUS
DEFINITION
pac82-164_8392.y2 pac82-164 Pseudomonas aeruginosa genomic clone
ACCESSION
BZ569481
VERSION
BZ569481.1 GI:27204170
KEYWORDS
GSS.

SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.

REFERENCE
1 (bases 1 to 945)
Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.

TITLE
Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

CONTACT: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers

1. .945

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="2-164"

/db_xref="taxon:287"

/clone="pac82-164_8392"

/clone_lib="pac82-164"

/notes="clinical isolate 2-164 Whole genomic shotgun library."

ORIGIN

Query Match 4.6%; Score 54.6; DB 8; Length 945;
Best Local Similarity 49.3%; Pred. No. 1.9e-05;
Matches 141; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 237 GCGGTTCTGTATATCAACCGGTACTCAATATCTTCTGGCGTGGTGCCATCATC 296

Db 778 GAGCATCTTCGATCGCAAAATCAGCTCTGCAACGCTCGNCAGCAGCTCCCATCTTC 719

QY 297 CGGAAGCGTTCCCGTCGGGATAAAAAATCGCGAGTGGCGGCTCCATGCAGACATCC 356

Db 718 TTGGCGCACTCTGGCGACTCAATAAGACGCTGATGGCATCCCGCAGAGCAGAGATCA 659

QY 357 CCCACGGGTAAAGCGTCCCTGTGCATCTTCTGAATGATCAGGGATCCCGCCGCTC 416

Db 658 CGCAGCGGCAGCAGCAGCGCCGTAACATCAGCTGGATTCATCCCGGCAACCCGGTACA 599

QY 417 TCACTGCGGATACGGGCGACCGCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCT 476

Db 598 TCGGTGTCACACCGCAGCGCCCGCAGCGCTTCGACCAGGACCTTGGGAAGCCCT 539

QY 477 TCATTTTCCGAGGCAATGACCACCACTGGCAATCCGGTAGACCG 522
Db 538 TCAGATAGAGGCGAGACGACTATATGCGAACGGCGGAGAGACAG 493

RESULT 3

BZ569276/c

LOCUS

DEFINITION

pac82-164_8326.x1 pac82-164 Pseudomonas aeruginosa genomic clone

ACCESSION

BZ569276

VERSION

BZ569276.1 GI:27203626

KEYWORDS

GSS.

SOURCE

Pseudomonas aeruginosa

ORGANISM

Pseudomonas aeruginosa

Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

Pseudomonadaceae; Pseudomonas.

REFERENCE

1 (bases 1 to 903)

Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,

Burns,J.L., Kaul,R. and Olsen,M.V.

Whole-Genome-Sequence variation among multiple isolates of

Pseudomonas aeruginosa library

J. Bacteriol. (2002) In press

CONTACT: Chris K. Raymond

Genome Center

University of Washington

Box 352145, Seattle, WA 98105-2145, USA

Tel: 2062216954

Fax: 2066857244

Email: craymond@u.washington.edu

Class: shotgun.

Location/Qualifiers

1. .903

/organism="Pseudomonas aeruginosa"

/mol_type="genomic DNA"

/strain="2-164"

/db_xref="taxon:287"

/clone="pac82-164_8326"

/clone_lib="pac82-164"

/notes="clinical isolate 2-164 Whole genomic shotgun library."

Query Match 4.6%; Score 54; DB 8; Length 903;

Best Local Similarity 51.7%; Pred. No. 2.9e-05;

Matches 123; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 285 GTCGCATCATCCGGAAGCGTTCCGGTCGGGATAAAAAATCGGCGAGTGGCGGTCCTCAT 344

Db 465 GCTCCCATCTCTTGGCGCAACTCTGGCGACTCAATAAGACGCTGATGGCATCCGCCAGA 406

QY 345 GCAGACATATCCCGCAGGGTAACAGCGTCCCTGTGCATTTCTTGAATGATCAGGG 404

Db 405 GCAGCAGAAATCAGCACCGGCGCAGCAGCGCCCGTAACATCAGCTGGATTCATCCCGG 346

QY 405 ATCCCGCCGCTCTCACTGGCGGATAACGGCAGCGCGGAGACTGACGCTTCAGCCAGTACC 464

Db 345 GAACCCGGTACATCGGTCTGCACACCGCAGCCCGCAGCGGACGCTTCGACCAGACC 286

QY 465 ATACCAAAACGTTCAATTTTCCGAAGGCATGACCAACACACTGGCAATCCGGTAGACCG 522

Db 285 TTGGGAAGCCCTTCACGATAAGAAGGCAGACGACTATATGCGAACGGCGGAAGACAG 228

RESULT 4

BZ569436/c

LOCUS

DEFINITION

pac82-164_8379.x1 pac82-164 Pseudomonas aeruginosa genomic clone

ACCESSION

BZ569436

VERSION

BZ569436.1 GI:27204054

KEYWORDS

GSS.

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	
Pseudomonas aeruginosa	Pseudomonas aeruginosa	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas.	1 (bases 1 to 920)	Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.I., Kaul,R. and Olsen,M.V.	Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library	J. Bacteriol. (2002) In press	
Genome Center	University of Washington	Box 352145, Seattle, WA 98105-2145, USA	Tel: 2062216954	Fax: 2066857244	Email: craymond@u.washington.edu	Class: shotgun.	
Location/Qualifiers	1. 920	/organism="Pseudomonas aeruginosa"	/mol_type="genomic DNA"	/strain="2-164"	/db_xref="taxon:287"	/clone_lib="pac2-164_8379"	/note="clinical isolate 2-164 Whole genomic shotgun library."
Query Match	4.5%;	Score 53;	DB 8;	Length 920;			
Best Local Similarity	51.3%;	Pred. No. 6.1e-05;					
Matches 122;	Conservative 0;	Mismatches 116;	Indels 0;	Gaps 0;			
Qy	285	GCTGCCATCATCCGGAAGCGTTCGGTCCGGATGAAATAAATCGCGCAGTCGCGCGTCCAT	344				
Db	776	GCTCCCACTCTTTGCCCACTCTGGCGACTCAATAAGAGCTGAATGGCATGCCCGAGA	717				
Qy	345	GCAGACACATCCCCACGGGTAAACAGCGTCCCTGTCTCATTTCTCAATGACATCAGGG	404				
Db	716	GCAGCAGATCATCCGACCGGACGAGCGCCGTATCATCAGCTGGATTGATCCCGG	657				
Qy	405	ATCCGCGCCGTCTCATCTGGCGATTAACCGGCACGCGCGAGACTGACGTTTCAGCCAGTACC	464				
Db	656	CAACCCCGGTACATCGTGTCTGTCACAAACCGCAGCCCGCAGCTCGACCGAGACC	597				
Qy	465	ATACCAACCGTTTCATTTTCGAGGATGACACACATGCGCAATCCGCTAGACCG	522				
Db	596	TTGCGAAGCCCTTCACGATGAAGGAGACGACTATATCCGAACGGGCGAAGACAG	539				
RESULT 5	BZ569367						
LOCUS	BZ569367	890 bp	DNA	linear	GSS 17-DEC-2002		
DEFINITION	pacs2-164_8360.x1 pacs2-164 Pseudomonas aeruginosa genomic clone						
ACCESSION	BZ569367						
VERSION	BZ569367.1	GI:27203887					
KEYWORDS	GSS.						
SOURCE	Pseudomonas aeruginosa						
ORGANISM	Pseudomonas aeruginosa						
REFERENCE	Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;						
AUTHORS	1 (bases 1 to 890)						
TITLE	Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M., Burns,J.I., Kaul,R. and Olsen,M.V.						
JOURNAL	Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library						
COMMENT	J. Bacteriol. (2002) In press						
	Contact: Chris K. Raymond						
	Genome Center						
	University of Washington						
	Box 352145, Seattle, WA 98105-2145, USA						

Seq primer: M13 Forward
Class: shotgun.

FEATURES

source
1. .755
Location/Qualifiers
/organism="Photorhabdus luminescens"
/mol_type="genomic DNA"
/strain="W14"
/db_xref="taxon:29488"
/clone="PLG00285"
/dev_stage="primary phase variant"
/clone_lib="Photorhabdus luminescens strain W14 M13 library"
/notes="Genomic DNA from strain W14 was size selected (1-2 kb) and then cloned into M13 Janus."

ORIGIN

Query Match 4.1%; Score 48.8; DB 8; Length 755;
Best Local Similarity 54.5%; Pred. No. 0.0013;
Matches 120; Conservative 0; Mismatches 97; Indels 3; Gaps 1;
QY 264 TCAATATCTTCTCTGGCGCTGCGCATCATCGGAAGCGTTCCGGTCGGGATAAAAA 323
Db 53 TCAATCGTTGACGAGCGGCTTTCCCATTTTTCCTCTCTGCTGTGTTGAGACATAAA 112
QY 324 TCGCGCAGTCGGCGTCCATGACAGACATCCCCACGGGTAAACGCGTCCCTGTCA 383
Db 113 AATTAACCGCATCGACAAATTGCTGMAACATCTCTGGCGTAAACCAAAATACCTGAATTA 172
QY 384 TTCTTCGATGACA---TCAGGATCCCGCCGCTCTCACTGGCGATAAGGCGACGCCG 440
Db 173 CTTTCATTACCGACAATTCCTGGTATGCTCCCATATTAAGTGGCAATAACCGCGCTGCCA 232
QY 441 GAGACTGACGCTTCAGCCAGTACCATACCAAAAGCTTCAT 480
Db 233 CAGCCATCGCTTCGCAATCGTAATCCCAACGCTTCGT 272.

RESULT 7

CD933246/c
LOCUS
DEFINITION
GR45.120F20F010719 GR45 Triticum aestivum cDNA clone GR45120P20,
mRNA sequence.
ACCESSION
CD933246
VERSION
CD933246.1 GI:32781010
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 608)
AUTHORS
Genoplante.
TITLE
Genoplante, a major partnership french program in plant genomics
JOURNAL
Unpublished (2003)
COMMENT
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).

FEATURES

source
1. .608
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="GR45120P20"
/tissue_type="grain (45 degrees per day after
pollination)";
/clone_lib="GR45"

ORIGIN

Query Match 4.1%; Score 48.2; DB 6; Length 608;
Best Local Similarity 56.7%; Pred. No. 0.0019;
Matches 89; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 371 CGTCCCTGTCAATTTCTTGAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAAC 430
Db 409 CTTTCCTTCTGATCTCAGGTATAATGTGAGTATGCTCCAGCTCGAGACGCGACCAC 350
QY 431 GGGCAGCGCGAGACTGACGCTTCAGCCAGTACCATACCAAGCTTCATTTTCCGAAGG 490
Db 349 CGGCACCCCGGATGACATTTGCCCTCCAAACAAACCCAGCGTCTCTGACTCAGAAGG 290
QY 491 CATGACCACACACTGGCAATCCGCTAGACCGGTAAC 527
Db 289 CATCACAAACACATCTCCACTGGCATAGGCTGTGAC 253
RESULT 8
CD909568/c
LOCUS
DEFINITION
G468.112P08F010820 G468 Triticum aestivum cDNA clone G468112P08,
mRNA sequence.
ACCESSION
CD909568
VERSION
CD909568.1 GI:32683892
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 648)
AUTHORS
Genoplante.
TITLE
Genoplante, a major partnership french program in plant genomics
JOURNAL
Unpublished (2003)
COMMENT
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (<http://www.genoplante.com>
and <http://genoplante-info.infobiogen.fr>).

FEATURES

source
1. .648
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G468112P08"
/tissue_type="grain (468 degrees per day after
pollination)";
/clone_lib="G468"

ORIGIN

Query Match 4.1%; Score 48.2; DB 6; Length 648;
Best Local Similarity 56.7%; Pred. No. 0.0019;
Matches 89; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
QY 371 CGTCCCTGTCAATTTCTTGAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAAC 430
Db 258 CTTTCCTTCTGATCTCAGGTATAATGTGAGTATGCTCCAGCTCGAGACGCGACCAC 199
QY 431 GGGCAGCGCGAGACTGACGCTTCAGCCAGTACCATACCAAGCTTCATTTTCCGAAGG 490
Db 198 CGGCACCCCGGATGACATTTGCCCTCCAAACAAACCCAGCGTCTCTGACTCAGAAGG 139
QY 491 CATGACCACACACTGGCAATCCGCTAGACCGGTAAC 527
Db 138 CATCACAAACACATCTCCACTGGCATAGGCTGTGAC 102

RESULT 9

CK153896	887 bp	mRNA	linear	EST 05-DEC-2003					
LOCUS	FGAS032577	Triticum aestivum FGAS: Talt2	Triticum aestivum cDNA,						
DEFINITION	mRNA sequence.								
ACCESSION	CK153896								
VERSION	CK153896.1	GI:38974450							
KEYWORDS	EST.								
SOURCE	Triticum aestivum (bread wheat)								
ORGANISM	Triticum aestivum								
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.								
AUTHORS	1 (bases 1 to 887)								
	Allard, P., Crosby, M.L., Danyluk, J., Eudes, P., Frick, M., Gaudet, D., Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Larocque, A., Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D., Pennak, C., Roach, J.L. and Sarhan, F.								
	Unpublished (2003)								
TITLE	Functional Genomics of Abiotic Stresses In Wheat and Canola Crops								
JOURNAL	Unpublished (2003)								
COMMENT	Contact: Wm L Crosby Bioinformatics University of Saskatchewan, Department of Computer Science 1C101 Engineering Building, 57 Campus Drive, Saskatoon, Saskatchewan, S7N 5A9, Canada Tel: 306 966 1769 Fax: 306 966 2033 Email: fgas.ests@cs.usask.ca This sequence is the direct result of the Base calling software Phred (default parameters). It is the raw base calls. To aid in the identification of the high quality insert the software Lucy (default parameters) has been run on this sequence. Lucy identified the region [122,762]. Plate: Talt242 row: J column: 03.								

```

FEATURES
source
Location/Qualifiers
1. .887
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Wheat line CI 14106"
/db_xref="taxon:4565"
/lab_host="DH5 alpha"
/clone_lib="Triticum aestivum FGAS: TaLtc2"
/note="Organ: Crown; Vector: pGEN-T; SSH (suppression subtractive hybridization) cDNA library from genotype CI14106 cold hardened at 2 C for 1 day (24 H) (tester) and subtracted against genotype Norstar cold hardened at 2 C for 21 days and 49 days (equal amount of cDNA pooled together before subtraction, driver). Nitro-pyrole anchored oligo-dT priming and non-directional cloning."

```

ORIGIN	Query Match	4.1%	Score 48.2;	DB 7;	Length 887;
	Best Local Similarity	56.7%;	Pred. No. 0.0021;		
	Matches 89;	Conservative 0;	Mismatches 68;	Indels 0;	Gaps 0;
Qy	371	CGTCCCTGTACATTTCTTGATGACATCAGGAGTCCGCGCGTCTCCTGCGGGATAAC	430		
Db	630	CTTTCTCTCTGTATCTCTCAGGTATAATGTGAGGTATGCTTCCAGCTCGAGCAGCGCCAC	689		
Qy	431	GGGCACGCCGGAGACTACGCTTCAGCGCAGTAGTACCATAACCAACCGCTTCATTTCCGAGG	490		
Db	690	CGGCACCCGGATGACATTGCCTCCACACACACAAACCACCGCTCTGTACTCAGAGG	749		
Qy	491	CATGACCACCACTATGGCAATCCGGTAGACCGGTAAAC	527		
Db	750	CATCAGAAACACATCTCCACTGGCATAGGCTGTGAC	786		

RESULT 10	CL982059/c	CL982059	1872 bp	DNA	linear	GSS 21-SEP-2004
LOCUS						
DEFINITION	OaIPSC046585 Oryza sativa Express Library Oryza sativa (indica cultivar-group) genomic, genomic survey sequence.					
ACCESSION	CL982059					

VERSION	CL982059.1	GI:52418604
KEYWORDS	GSS.	
SOURCE	Oryza sativa (indica cultivar-group)	
ORGANISM	Oryza sativa (indica cultivar-group)	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.	
REFERENCE	1 (bases 1 to 1872)	
AUTHORS	Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M., Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L., Wong, G.K.S., Deng, X.W. and Wang, J.	
TITLE	An analysis of transcriptional regulation of the rice genome and its comparison to Arabidopsis	
JOURNAL	Unpublished (2004)	
COMMENT	Contact: Chen Chen Department of Bioinformatic Beijing Institute of Genomics Chinese Academy of Sciences, Beijing 101300, China Tel: 86-10-80481559 Fax: 86-10-80488676 Email: chenchen@genomics.org.cn Rice genomic sequence. Class: exon-trapped. Location/Qualifiers 1. .1872	
FEATURES	/organism="Oryza sativa (indica cultivar-group)" /mol_type="genomic DNA" /db_xref="taxon:39946" /clone_lib="Oryza sativa Express Library" /note="Oryza sativa exon trapped genomic sequences"	
ORIGIN		

Query Match	4.0%;	Score	46.8;	DB	9;	Length	1872;
Best Local Similarity	45.8%;	Pred. No.	0.0077;				
Matches	162;	Conservative	0;	Mismatches	192;	Indels	0;
Gaps	0;						
Qy	285	GCTGCCATCATCGGAAGCGTTC	CGGT	CGGATAAAAAATCGCG	CACTCGCGCGTCCAT	344	
Db	1790	GCTTCGCCATGAGAGATCC	AGCGCT	GCGCGGAGCGCCGAT	TGGCTTCGGTGATG	1731	
Qy	345	GCACAGACATCCCCACGGGTAA	CAGCGTCCCTGTGCACAT	TTTCTGAAATGACATCAGGG	404		
Db	1730	GCGGCGACGTCGAGCGCTCGGT	CACGAAGCCGTTCTGGCCGCGGTGATAAACTCGGCG	1671			
Qy	405	ATCCGCGCGCTCTCACTGGGATA	ACGGGACAACGCGGAGACTACGCTT	CAGCCAGTACC	464		
Db	1670	CCGCCCGAGTGGTGTGTGTGATA	ACCGGAGCGCCGAGGACATCGCTT	CCAGAATCACA	1611		
Qy	465	ATACCAAGCGTTTCATTTTTCGA	AGGCGATGACCACACACTGCGCAAT	CCCGGTAGACCGGT	524		
Db	1610	TTCCGGAGCGATCGTACAGCGT	CGGAGCGAGCGGCATCCGGCGCTGAT	TAGACGGC	1551		
Qy	525	AACGCTGGGAAAAGGCGACCTG	CGCATTAACACATCTCCGCTCAT	TTCCAGGTTGTTCTGTG	584		
Db	1550	AGAGTCTGCTTCTGCAGGCCCA	TAAAGTGAATCGGATCGCGCAGCC	CCGCGACTGCGCC	1491		
Qy	585	TGCTGACGACAGCGTGTCTGTA	TTCTTACGCCCGCGGCCAC	ACGAGCCAG	638		
Db	1490	AGCGCCCGGTAAACGCTTTTTCGG	CTTGTCTTTTACCGACACACCAAGAGTGGCTG	1437			

RESULT 11	CA735693	328 bp	linear	EST 26-NOV-2002
LOCUS	wpils.pk005.g7			
DEFINITION	Triticum aestivum cDNA clone wpils.pk005.g7 5' end, mRNA sequence.			
ACCESSION	CA735693			
VERSION	CA735693.1	GI:25551291		
KEYWORDS	EST.			
SOURCE	Triticum aestivum (bread wheat)			
ORGANISM	Triticum aestivum			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			

```
REFERENCE
AUTHORS
Tingey,S.V., Powell,W., Walters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafey,M.K.
TITLE
DuPont Wheat cDNA Sequence
JOURNAL
Unpublished (2002)
COMMENT
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
FEATURES
source
Location/Qualifiers
1..328
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wpils.pk005.g7"
/tissue_type="leaf"
/lab_host="DH10B"
/clone_lib="wpils"
/notes="vector: pGEM-T Easy; Site_1: SmaI; Wheat, Polk
cultivar (resistant), infected with septoria tritici
strain A 24 hours after infection"
ORIGIN
Query Match 3.9%; Score 46.6; DB 6; Length 328;
Best Local Similarity 56.1%; Pred. No. 0.005;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430
Db 32 CTTTCTTCTGATCCTCAGTATATGTGATGCTCCAGCTCGACGAGCCAC 91
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCCATACATACCAACGCTTCATTTCCGAAGG 490
Db 92 CGGAGCTCCGGATGACATTCCTCCCAACACAAACCCAGGCTCTCGGACTCAGAAGG 151
QY 491 CATGACACACACTGCGCAATCCGGTAGACCGGTAAC 527
Db 152 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 188
RESULT 12
CD881543
LOCUS
F1.103123R010627 F1 Triticum aestivum cDNA clone F1103123, mRNA
DEFINITION
sequence.
ACCESSION
CD881543.1 GI:32641069
VERSION
CD881543
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 507)
Genoplante.
Genoplante, a major partnership french program in plant genomics
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
source
Location/Qualifiers
1..507
/organism="Triticum aestivum"
```

```
REFERENCE
AUTHORS
Tingey,S.V., Powell,W., Walters,P., Dolan,M., Hainey,C., Yuan,Z.,
Miao,G., Caraher,N. and Hanafey,M.K.
TITLE
DuPont Wheat cDNA Sequence
JOURNAL
Unpublished (2002)
COMMENT
Contact: Scott V. Tingey
Crop Genetics
E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Tel: 302-631-2602
Fax: 302-631-2607
Email: Scott.V.Tingey@USA.dupont.com
Seq primer: M13.
FEATURES
source
Location/Qualifiers
1..328
/organism="Triticum aestivum"
/mol_type="mRNA"
/db_xref="taxon:4565"
/clone="wpils.pk005.g7"
/tissue_type="leaf"
/lab_host="DH10B"
/clone_lib="wpils"
/notes="vector: pGEM-T Easy; Site_1: SmaI; Wheat, Polk
cultivar (resistant), infected with septoria tritici
strain A 24 hours after infection"
ORIGIN
Query Match 3.9%; Score 46.6; DB 6; Length 328;
Best Local Similarity 56.1%; Pred. No. 0.005;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430
Db 32 CTTTCTTCTGATCCTCAGTATATGTGATGCTCCAGCTCGACGAGCCAC 91
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCCATACATACCAACGCTTCATTTCCGAAGG 490
Db 92 CGGAGCTCCGGATGACATTCCTCCCAACACAAACCCAGGCTCTCGGACTCAGAAGG 151
QY 491 CATGACACACACTGCGCAATCCGGTAGACCGGTAAC 527
Db 152 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 188
RESULT 13
BJ544067/c
LOCUS
BJ544067 K. Sato unpublished cDNA library, cv. Akashinriki
DEFINITION
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak18m01 5', mRNA sequence.
ACCESSION
BJ544067.1 GI:24962057
VERSION
BJ544067
KEYWORDS
EST.
SOURCE
Hordeum vulgare subsp. vulgare
ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
1 (bases 1 to 569)
Sato,K., Saisho,D. and Takeda,K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
Location/Qualifiers
1..569
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baak18m01"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"
ORIGIN
Query Match 3.9%; Score 46.6; DB 4; Length 569;
Best Local Similarity 56.1%; Pred. No. 0.006;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATTTCTTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430
Db 451 CTTTCTTCTGCTCTCAGTATATGTGATGCTCCAGCTCGACGAGCCAC 392
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCCAGTACCAACGCTTCATTTCCGAAGG 490
```

```
Db 391 CGGAATCCCGATGACATGCTCCAAACACACAAACCCAGCGCTCTGATTGAGAAGG 332
Qy 491 CATGACACACCACTGGCAATCCGGTAGACCGGTAAC 527
Db 331 CATCAAAACACATCCCGCTGATAGGCTGTGAC 295

RESULT 14
BJ551686
LOCUS BJ551686 609 bp mRNA linear EST 14-NOV-2002
DEFINITION BJ551686 K. Sato unpublished cDNA library, strain H602 adult,
heading stage top three leaves Hordeum vulgare subsp. spontaneum
cDNA clone bah35c20 3', mRNA sequence.
ACCESSION BJ551686
VERSION BJ551686.1 GI:24970137
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. spontaneum
ORGANISM Hordeum vulgare subsp. spontaneum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 609)
AUTHORS Sato, K., Saisho, D. and Takeda, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..609
/organism="Hordeum vulgare subsp. spontaneum"
/mol_type="mRNA"
/strain="H602"
/sub_species="spontaneum"
/db_xref="taxon:77009"
/clone="bah35c20"
/tissue_type="top three leaves"
/dev_stage="adult, heading stage"
/clone_lib="K. Sato unpublished cDNA library, strain H602
adult, heading stage top three leaves"
ORIGIN
Query Match 3.9%; Score 46.6; DB 4; Length 609;
Best Local Similarity 56.1%; Pred. No. 0.0061;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
Qy 371 CGTCCCTGTGACATTTCTTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430
Db 423 CTTTCCTTCTGGTCTCTCAGGATATGTCAGTATGCTCCAGCTCGAGCGGACAC 482
Qy 431 GGGCAGCCCGGAGACTGACGCTTCAGCGATGATACATACCAAGCGTTTCATTTCCGAAGG 490
Db 483 CGGAATCCCGATGACATGCTCCAAACACAAACCCAGCGCTCTGATTGAGAAGG 542
Qy 491 CATGACACACCACTGGCAATCCGGTAGACCGGTAAC 527
Db 543 CATCAAAACACATCCCGCTGATAGGCTGTGAC 579

RESULT 15
AV909088/c
LOCUS AV909088 615 bp mRNA linear EST 18-JAN-2002
DEFINITION AV909088 K. Sato unpublished cDNA library, cv. Akashinriki
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak11i22 5', mRNA sequence.
ACCESSION AV909088
VERSION AV909088.1 GI:18204518
KEYWORDS EST.
```

```
SOURCE Hordeum vulgare subsp. vulgare
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 615)
AUTHORS Sato, K., Saisho, D. and Takeda, K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..615
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baak11i22"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"
ORIGIN
Query Match 3.9%; Score 46.6; DB 1; Length 615;
Best Local Similarity 56.1%; Pred. No. 0.0062;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
Qy 371 CGTCCCTGTGACATTTCTTGAATGACATCAGGATCCCGCGCTCTCACTGGCGATAAC 430
Db 444 CTTTCCTTCTGGTCTCTCAGGATATGTCAGTATGCTCCAGCTCGAGCGGACAC 385
Qy 431 GGGCAGCCCGGAGACTGACGCTTCAGCGATGATACATACCAAGCGTTTCATTTCCGAAGG 490
Db 384 CGGAATCCCGATGACATGCTCCAAACACAAACCCAGCGCTCTGATTGAGAAGG 325
Qy 491 CATGACACACCACTGGCAATCCGGTAGACCGGTAAC 527
Db 324 CATCAAAACACATCCCGCTGATAGGCTGTGAC 288

RESULT 16
BJ264853
LOCUS BJ264853 629 bp mRNA linear EST 08-APR-2002
DEFINITION BJ264853 Y. Ogihara unpublished cDNA library, Wh_h Triticum
aestivum cDNA clone whh9f24 3', mRNA sequence.
ACCESSION BJ264853
VERSION BJ264853.1 GI:20085028
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE 1 (bases 1 to 629)
AUTHORS Ogihara, Y. and Murai, K.
TITLE Expressed genes in Triticum aestivum
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source
1..629
```

```
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="whh9f24"
/tissue_type="spike at heading date"
/dev_stage="Feekes", scale 10.5"
/clone_lib="Y. Ogiwara unpublished cDNA library, wh_h"

ORIGIN

Query Match          3.9%; Score 46.6; DB 4; Length 629;
Best Local Similarity 56.1%; Pred. No. 0.0062;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 371 CGTCCCTGTGCACATTTCTTCTGAATGACATCAGGATCCCGCGTCTCTCACTGGCGATAAC 430
Db 329 CTTTCTTCTGATCCTCAGTATAATGTTCAGGTATGCTTCCAGCTCGAGCAGGCCAC 388
QY 431 GGGCAGCCCGAGACTGACGCTTTCAGCCAGTACCATCAACAAAGCTTTCATTTTCCGAAGG 490
Db 389 CGGAGCTCCGGATGACATTGCTCTCAACACAAACAAACCCAGCGTCTCGGACTCAGAAGG 448
QY 491 CATGACACCACTGCGCAATCCGGTAGACCGGTAAC 527
Db 449 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 485

RESULT 17
BJ546361
LOCUS
DEFINITION
BJ546361 K. Sato unpublished cDNA library, cv. Akashinriki
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak39j08 3', mRNA sequence.
BJ546361
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 640)
Sato, K., Saisho, D. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .640
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baak39j08"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"

ORIGIN

Query Match          3.9%; Score 46.6; DB 4; Length 640;
Best Local Similarity 56.1%; Pred. No. 0.0062;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 371 CGTCCCTGTGCACATTTCTTCTGAATGACATCAGGATCCCGCGTCTCTCACTGGCGATAAC 430
Db 484 CTTTCTTCTGCTCCTCAGGAATAATGTTCAGGTATGCTTCCAGCTCGAGCAGGCCAC 543
```

```
QY 431 GGGCAGCCCGAGACTGACGCTTTCAGCCAGTACATCAACAAAGCTTTCATTTTCCGAAGG 490
Db 544 CGGAATCCCGATGACATTCCTTCCACACAAACAAACCCAGCGTCTCTGATTCAGAAGG 603
QY 491 CATGACACCACTGCGCAATCCGGTAGACCGGTAAC 527
Db 604 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 640

RESULT 18
BJ545036/c
LOCUS
DEFINITION
BJ545036 K. Sato unpublished cDNA library, cv. Akashinriki
vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
baak42j16 5', mRNA sequence.
BJ545036
VERSION
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 641)
Sato, K., Saisho, D. and Takeda, K.
Barley EST sequencing project in NIG and Okayama Univ
Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .641
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baak42j16"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cDNA library, cv.
Akashinriki vegetative stage leaves"

ORIGIN

Query Match          3.9%; Score 46.6; DB 4; Length 641;
Best Local Similarity 56.1%; Pred. No. 0.0062;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 371 CGTCCCTGTGCACATTTCTTCTGAATGACATCAGGATCCCGCGTCTCTCACTGGCGATAAC 430
Db 641 CTTTCTTCTGCTCCTCAGGAATAATGTTCAGGTATGCTTCCAGCTCGAGCAGGCCAC 582
QY 431 GGGCAGCCCGAGACTGACGCTTTCAGCCAGTACCATCAACAAAGCTTTCATTTTCCGAAGG 490
Db 581 CGGAATCCCGATGACATTTGCTTCCAAACAAACCCAGCGTCTCTGATTCAGAAGG 522
QY 491 CATGACACCACTGCGCAATCCGGTAGACCGGTAAC 527
Db 521 CATCAAAACACATCCCACTGGCATAGGCGCTGTGAC 485

RESULT 19
BJ552866
LOCUS
DEFINITION
BJ552866 K. Sato unpublished cDNA library, strain H602 adult,
heading stage top three leaves Hordeum vulgare subsp. spontaneum
cDNA clone bah32f09 3', mRNA sequence.
BJ552866
ACCESSION
```

```

VERSION      BJ52866.1  GI:24971317
KEYWORDS     EST.
SOURCE       Hordeum vulgare subsp. spontaneum
ORGANISM     Hordeum vulgare subsp. spontaneum
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
             Poideae; Triticeae; Hordeum.
REFERENCE    1 (bases 1 to 647)
AUTHORS      Sato, K., Saisho, D. and Takeda, K.
TITLE        Barley EST sequencing project in NIG and Okayama Univ
JOURNAL      Unpublished (2002)
COMMENT      Contact: Tadasu Shin-i
             Center For Genetic Resource Information
             National Institute of Genetics
             1111 Yata, Mishima, Shizuoka 411-8540, Japan
             Tel: 81-559-81-6856
             Fax: 81-559-81-6855
             Email: tshini@genes.nig.ac.jp.

FEATURES     Location/Qualifiers
             1..647
             /organism="Hordeum vulgare subsp. spontaneum"
             /mol_type="mRNA"
             /strain="H602"
             /sub_species="spontaneum"
             /db_xref="taxon:77009"
             /clone="bah32f09"
             /tissue_type="top three leaves"
             /dev_stage="adult, heading stage"
             /clone_lib="K. Sato unpublished cdna library, strain H602
             adult, heading stage top three leaves"

ORIGIN
Query Match      3.9%; Score 46.6; DB 4; Length 647;
Best Local Similarity 56.1%; Pred. No. 0.0063;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy  371  CGTCCCTGTCACATCTCTCAATGACATCAGGATCCGCCGCTCTCACTGGCGATAAC 430
Db  460  CTTTCTCTCTCGTCTCAGGAATAATGTGATGATGCTCCAGCTCGAGCGACCCAC 519
Qy  431  GGGCAGCGCGAGACTGACGCTTCAGCGAGTACCATAACCAACGCTTCATTTCCGAAGG 490
Db  520  CGGAATCTCCGATGACATGCTCTCAACACACAAACCCAGCGTCTCTGATTCAGAAGG 579
Qy  491  CATGACCACACATCGGCAATCCGGTAGACCGGTAAAC 527
Db  580  CATCAACAACATCCCCACTGGCATAGGCGCTGTGAC 616

RESULT 20
BJ546519
LOCUS      BJ546519 651 bp mRNA linear EST 14-NOV-2002
DEFINITION vegetative stage leaves Hordeum vulgare subsp. vulgare cdna clone
            baak42j16 3', mRNA sequence.
ACCESSION BJ546519
VERSION    BJ546519.1  GI:24964935
KEYWORDS   EST.
SOURCE     Hordeum vulgare subsp. vulgare
ORGANISM   Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 651)
AUTHORS    Sato, K., Saisho, D. and Takeda, K.
TITLE      Barley EST sequencing project in NIG and Okayama Univ
JOURNAL    Unpublished (2002)
COMMENT    Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855

```

```

Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..651
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Akashinriki"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="baak42j16"
/tissue_type="leaves"
/dev_stage="vegetative stage"
/clone_lib="K. Sato unpublished cdna library, cv.
Akashinriki vegetative stage leaves"

ORIGIN
Query Match      3.9%; Score 46.6; DB 4; Length 651;
Best Local Similarity 56.1%; Pred. No. 0.0063;
Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy  371  CGTCCCTGTCACATCTCTCAATGACATCAGGATCCGCCGCTCTCACTGGCGATAAC 430
Db  443  CTTTCTCTCTCGTCTCAGGAATAATGTGATGATGCTCCAGCTCGAGCGACCCAC 502
Qy  431  GGGCAGCGCGAGACTGACGCTTCAGCGAGTACCATAACCAACGCTTCATTTCCGAAGG 490
Db  503  CGGAATCTCCGATGACATGCTCTCAACACACAAACCCAGCGTCTCTGATTCAGAAGG 562
Qy  491  CATGACCACACATCGGCAATCCGGTAGACCGGTAAAC 527
Db  563  CATCAACAACATCCCCACTGGCATAGGCGCTGTGAC 599

RESULT 21
AV910971
LOCUS      AV910971 669 bp mRNA linear EST 18-JAN-2002
DEFINITION vegetative stage leaves Hordeum vulgare subsp. vulgare cdna clone
            baak11122 3', mRNA sequence.
ACCESSION AV910971
VERSION    AV910971.1  GI:18206747
KEYWORDS   EST.
SOURCE     Hordeum vulgare subsp. vulgare
ORGANISM   Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Poideae; Triticeae; Hordeum.
REFERENCE  1 (bases 1 to 669)
AUTHORS    Sato, K., Saisho, D. and Takeda, K.
TITLE      Barley EST sequencing project in NIG and Okayama Univ
JOURNAL    Unpublished (2002)
COMMENT    Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.

FEATURES     Location/Qualifiers
             1..669
             /organism="Hordeum vulgare subsp. vulgare"
             /mol_type="mRNA"
             /cultivar="Akashinriki"
             /sub_species="vulgare"
             /db_xref="taxon:112509"
             /clone="baak11122"
             /tissue_type="leaves"
             /dev_stage="vegetative stage"
             /clone_lib="K. Sato unpublished cdna library, cv.
             Akashinriki vegetative stage leaves"

ORIGIN
Query Match      3.9%; Score 46.6; DB 1; Length 669;
Best Local Similarity 56.1%; Pred. No. 0.0063;

```

Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 371 CGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAAC 430
 Db 460 CTTTCTCTCTGCTCTCAGGAATAATGTAGGTATGCTTCCAGCTCGAGCGGACCAC 519
 Qy 431 GGGCAGCCCGAGACTGACGCTTCAGCCAGTACCATCAAAACGCTTTCATTTCCGAAGG 490
 Db 520 CGGAACCTCCCGATGACATTGCTCCAAACACAAACCCAGCGCTCTCTGATTGAGAAGG 579
 Qy 491 CATGACCAACACATCGGCATCCGGTAGACCGGTAAC 527
 Db 580 CATCAAAACACATCCCCACTGGCATAGGCTGTGAC 616

RESULT 22

LOCUS
 DEFINITION BJ551738 683 bp mRNA linear EST 14-NOV-2002
 heading stage top three leaves Hordeum vulgare subsp. spontaneum
 cDNA clone bah40116 3', mRNA sequence.

ACCESSION

VERSION BJ551738

KEYWORDS

SOURCE

ORGANISM

Hordeum vulgare subsp. spontaneum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

1 (bases 1 to 683)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..683

/organism="Hordeum vulgare subsp. spontaneum"

/mol_type="mRNA"

/strain="H602"

/sub_species="spontaneum"

/db_xref="taxon:77009"

/clone="bah40116"

/tissue_type="top three leaves"

/dev_stage="adult, heading stage"

/clone_lib="K. Sato unpublished cDNA library, strain H602"

adult, heading stage top three leaves"

Query Match 3.9%; Score 46.6; DB 4; Length 683;

Best Local Similarity 56.1%; Pred. No. 0.0064;

Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 371 CGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAAC 430

Db 473 CTTTCTCTCTGCTCTCAGGAATAATGTAGGTATGCTTCCAGCTCGAGCGGACCAC 532

Qy 431 GGGCAGCCCGAGACTGACGCTTCAGCCAGTACCATCAAAACGCTTTCATTTCCGAAGG 490

Db 533 CGGAACCTCCCGATGACATTGCTCCAAACACAAACCCAGCGCTCTCTGATTGAGAAGG 592

Qy 491 CATGACCAACACATCGGCATCCGGTAGACCGGTAAC 527

Db 593 CATCAAAACACATCCCCACTGGCATAGGCTGTGAC 629

RESULT 23

AV945422

LOCUS
 DEFINITION AV945422 K. Sato unpublished cDNA library, strain H602 adult,
 heading stage top three leaves Hordeum vulgare subsp. spontaneum
 cDNA clone bah261i5 3', mRNA sequence.

ACCESSION AV945422 GI:18241219

VERSION AV945422.1

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. spontaneum

ORGANISM Hordeum vulgare subsp. spontaneum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

1 (bases 1 to 693)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..693

/organism="Hordeum vulgare subsp. spontaneum"

/mol_type="mRNA"

/strain="H602"

/sub_species="spontaneum"

/db_xref="taxon:77009"

/clone="bah261i5"

/tissue_type="top three leaves"

/dev_stage="adult, heading stage"

/clone_lib="K. Sato unpublished cDNA library, strain H602"

adult, heading stage top three leaves"

Query Match 3.9%; Score 46.6; DB 1; Length 693;

Best Local Similarity 56.1%; Pred. No. 0.0064;

Matches 88; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 371 CGTCCCTGTACATCTTCTGAATGACATCAGGATCCCGCCGCTCTCACTGGCGATAAC 430

Db 487 CTTTCTCTCTGCTCTCAGGAATAATGTAGGTATGCTTCCAGCTCGAGCGGACCAC 546

Qy 431 GGGCAGCCCGAGACTGACGCTTCAGCCAGTACCATCAAAACGCTTTCATTTCCGAAGG 490

Db 547 CGGAACCTCCCGATGACATTGCTCCAAACACAAACCCAGCGCTCTCTGATTGAGAAGG 606

Qy 491 CATGACCAACACATCGGCATCCGGTAGACCGGTAAC 527

Db 607 CATCAAAACACATCCCCACTGGCATAGGCTGTGAC 643

RESULT 24

BJ552083

LOCUS

DEFINITION BJ552083 713 bp mRNA linear EST 14-NOV-2002

heading stage top three leaves Hordeum vulgare subsp. spontaneum

cDNA clone bah47d22 3', mRNA sequence.

ACCESSION BJ552083

VERSION BJ552083.1 GI:24970534

KEYWORDS EST.

SOURCE Hordeum vulgare subsp. spontaneum

ORGANISM Hordeum vulgare subsp. spontaneum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

1 (bases 1 to 713)

Sato, K., Saisho, D. and Takeda, K.

Barley EST sequencing project in NIG and Okayama Univ

Unpublished (2002)

Contact: Tadasu Shin-i

Query Match	3.9%	Score 45.8;	DB 4;	Length 647;
Best Local Similarity	56.2%;	Pred. NO. 0.011;		
Matches	86;	Conservative 0;	Mismatches 67;	Indels 0; Gaps 0;
QY	371	CGTCCCTGTGCACATTTCTCTGAATGCATCATCAGGAGATCCGGCCGGTCTCACTGCGGATAAC	430	
DB	458	CTTCCCTCCAGTCTCTCGGGATATATGTCTGGGGATCCGCGCGCGGACCGCCACCCAC	517	


```

SOURCE
ORGANISM
Oryza minuta
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 538)
Oryza minuta HybridZAP-2.1 XR library
Unpublished (2003)
Contact: Jeong Sheop Shin
Plant Molecular Genetics
Graduate School of Biotechnology, University of Korea
136-701 Anan-dong 5/1 Seoul, Korea
Tel: 00 82 2 3290 3430
Fax: 00 82 2 927 9028
Email: jshin@kucn.korea.ac.kr.
FEATURES
source
1..538
Location/Qualifiers
/organism="Oryza minuta"
/mol_type="mRNA"
/db_xref="taxon:63629"
/dev_stage="4-weeks after germination"
/clone_lib="Oryza minuta HybridZAP-2.1 XR library"
/note="Organ: immature leaf"
ORIGIN
Query Match 3.9%; Score 45.6; DB 6; Length 538;
Best Local Similarity 55.8%; Pred. No. 0.012;
Matches 87; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 371 CGTCCGTGCACATCTCTGATGACATCAGGATCCCGCGCTCTCATCTGGCGGATAAC 430
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
226 CTTACCTCTGATCCTCTGGTATAATATCAGGTATTCCTCCAGCGGAGCTGCAACTAC 167
QY 431 GGGCAGCGCGGAGACTGACCTTCAGCCAGTACCAACCAAGCTTCATTTCCGAAGG 490
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
166 GGGACTCTTGATGACATAGCTCCATACACAAACCAAGGCTTCTGACTCAGAGG 107
QY 491 CATTGACACACACTGGCAATCCGGTAGACCGGTAA 526
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 CATCAAAACATCCCACTGGCATAGGCGCTGTGA 71

RESULT 30
CF587963/c
LOCUS
DEFINITION
USDA-FP 121100-054 Acyrthosiphon pisum, Pea Aphid Acyrthosiphon
pisum cDNA clone WHAP-008_E07_5', mRNA sequence.
ACCESSION
CF587963
VERSION
CF587963.1 GI:35508191
KEYWORDS
EST.
SOURCE
Acyrthosiphon pisum (pea aphid)
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
1 (bases 1 to 313)
HUNTER, W.B., DANG, P.M., STERN, D. and LAPORTE, S.L.
Gene expression in Acyrthosiphon pisum (Hemiptera: Aphididae)
Unpublished (2004)
Contact: Wayne Hunter; Phat Dang
US. Horticultural Research Lab, Ft. Pierce, FL
USDA-ARS
2001 South Rock Rd., Ft. Pierce, FL 34945, USA
Tel: (772) 462-5898, (772) 462-5940
Fax: (772) 462-5986
Email: Whunter@ushrl.ars.usda.gov
Seq primer: T3 Primer.
Location/Qualifiers
1..313
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/db_xref="taxon:7029"
/clone_lib="WHAP035_A05"
/sex="Mixed population"
/tissue_type="whole body, nymphs and adults"
/dev_stage="All"
/dev_host="XLI-Blue"
/clone_lib="Acyrthosiphon pisum, Pea Aphid"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; Acyrthosiphon pisum colonies. Library by Srin-
Kambampati, Peter Dearden; David Stern, Department of
Ecology and Evolutionary Biology, Princeton University,
NJ. (609) 258-0759. A high quality EST with at least 100
contiguous bases at Trace Tuner score of 20 or better.
Sequencing of clones by Dr. PM Dang, US. Horticultural
Research Lab, Ft. Pierce, FL. (772) 462- 5940."
FEATURES
source
1..313
Location/Qualifiers
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/db_xref="taxon:7029"

```

```

/clone="WHAP-008_E07"
/sex="Mixed population"
/tissue_type="whole body, nymphs and adults"
/dev_stage="All"
/lab_host="XLI-Blue"
/clone_lib="Acyrthosiphon pisum, Pea Aphid"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; Acyrthosiphon pisum colonies. Library by Srin-
Kambampati, Peter Dearden; David Stern, Department of
Ecology and Evolutionary Biology, Princeton University,
NJ. (609) 258-0759. A high quality EST with at least 100
contiguous bases at Trace Tuner score of 20 or better.
Sequencing of clones by Dr. PM Dang, US. Horticultural
Research Lab, Ft. Pierce, FL. (772) 462- 5940."
ORIGIN
Query Match 3.8%; Score 45.4; DB 7; Length 313;
Best Local Similarity 63.1%; Pred. No. 0.012;
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
QY 11 TGGAAAAAGCCAAATAATAAATTGCCCATCCAGCGGCTCCAGCTGAAAGTAGGCC 70
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
169 TAGAGAAAAACCAAAAGTAGAGCAGTGCCCAACCCAGGGGCTTCCAGCTAAAAAGGTT 110
QY 71 TGTTCGTCCGGTATTAAATGCAATGACCGTCCCGCTATTATAACAATGT 121
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109 TAGCCTGGACTCTCTGTGGAAATGATAGTAAGAACCTGTCTTGAATAAATAT 59

RESULT 31
CN583481/c
LOCUS
DEFINITION
USDA-FP 126545 Acyrthosiphon pisum, Pea Aphid Acyrthosiphon pisum
cDNA clone WHAP035_A05, mRNA sequence.
ACCESSION
CN583481
VERSION
CN583481.1 GI:46995204
KEYWORDS
EST.
SOURCE
Acyrthosiphon pisum (pea aphid)
ORGANISM
Acyrthosiphon pisum
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
Aphidoidea; Aphididae; Macrosiphini; Acyrthosiphon.
1 (bases 1 to 313)
HUNTER, W.B., DANG, P.M., STERN, D. and LAPORTE, S.L.
Gene expression in Acyrthosiphon pisum (Hemiptera: Aphididae)
Unpublished (2004)
Contact: Wayne Hunter; Phat Dang
US. Horticultural Research Lab, Ft. Pierce, FL
USDA-ARS
2001 South Rock Rd., Ft. Pierce, FL 34945, USA
Tel: (772) 462-5898, (772) 462-5940
Fax: (772) 462-5986
Email: Whunter@ushrl.ars.usda.gov.
Location/Qualifiers
1..313
/organism="Acyrthosiphon pisum"
/mol_type="mRNA"
/db_xref="taxon:7029"
/clone_lib="WHAP035_A05"
/sex="Mixed population"
/tissue_type="whole body, nymphs and adults"
/dev_stage="All"
/dev_host="XLI-Blue"
/clone_lib="Acyrthosiphon pisum, Pea Aphid"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; Acyrthosiphon pisum colonies. Library by Srin-
Kambampati, Peter Dearden; David Stern, Department of
Ecology and Evolutionary Biology, Princeton University,
NJ. (609) 258-0759. A high quality EST with at least 100
contiguous bases at Trace Tuner score of 20 or better.
Sequencing of clones by Dr. PM Dang, US. Horticultural
Research Lab, Ft. Pierce, FL. (772) 462- 5940."
ORIGIN

```

```
Query Match          3.8%; Score 45.4; DB 7; Length 313;
Best Local Similarity 63.1%; Pred. No. 0.012;
Matches 70; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 11 TGAAGAAAAGCCAAATAAATAATTCCTCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 70
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 TAGAGAAAACCAAAAGTAGAGCAGTGCCTCCCAAGCCAGGGCTTTCCAGCTAAAAAGAGTT 110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 71 TGTTCGTCCGGTATTAAATGATTCACCGTCCCGCTATTATAACAATGT 121
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 TAGCCTGGACTTCTGTGGAATGATAGTAAGAACCTGTCTGTGAATAAATAT 59
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 32
B0763870/c
LOCUS
DEFINITION
    B0763870 495 bp mRNA linear EST 26-JUL-2002
    Hordeum vulgare subsp. vulgare cDNA clone EBro03_SQ005_K13 5', mRNA
    sequence.
ACCESSION
    B0763870
KEYWORDS
    EST.
SOURCE
    B0763870.1 GI:21972342
ORGANISM
    Hordeum vulgare subsp. vulgare
    Hordeum vulgare subsp. vulgare
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Poideae; Triticeae; Hordeum.
REFERENCE
    1 (bases 1 to 495)
AUTHORS
    Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
    Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
TITLE
    Development of Barley Transcriptome Resources
JOURNAL
    Unpublished (2001)
COMMENT
    Contact: Waugh R, Marshall DF
    Genome Dynamics/Computational Biology
    Scottish Crop Research Institute
    Invergowrie, Dundee, DD2 5DA, Scotland, UK
    Tel: 00 44 1382 562731
    Fax: 00 44 1382 562426
    Email: est@scri.sari.ac.uk.
FEATURES
    source
    1..495
        /organism="Hordeum vulgare subsp. vulgare"
        /mol_type="mRNA"
        /cultivar="Optic"
        /sub_species="vulgare"
        /db_xref="taxon:112509"
        /clone="EBro03_SQ005_K13"
        /tissue_type="root"
        /dev_stage="3 week"
        /lab_host="DH10B"
        /clone_lib="root, 3 week, waterlogged, cv Optic, EBro03"
        /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
        Non-normalised library, directionally cloned into pSPORT1.
        Derived from roots of 3 week old waterlogged barley
        plants. Developed as part of the barley transcriptome
        resources of BBSRC/SEERAD funded cereal IGF (Investigating
        Gene Function) project."
ORIGIN

Query Match          3.8%; Score 45; DB 5; Length 495;
Best Local Similarity 49.8%; Pred. No. 0.019;
Matches 114; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 404 GATCCGCCCGCTCTCACTGGCGATAACGGCAGCGCGGAGACTGACGCTTCAGCCAGTAC 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 261 GAAGTCGAGCGGACGCTGCGAGGAGCGCGACGCGCGGATGAGCGCAACGCGCGGGC 202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 464 CATACCAAAAGCGTTTCATTTTCCGAAGCATGACCACACATCGGCAATCGGTAGACCGG 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 CAGCGCAGACCCGATGGCGCGCGGGATGGACATGATCCCGCGCCCAACACACGCTCGT 142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 524 TAACGCTGGGAAAAGGGCACCTGTGCCATTAACACATCTCCGCTCATTTCCAGGTTCTGT 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 141 CGACACGTTTGAACACACCGCGCGGAGGACCGACGCGTTCGCGGAGGCGCCCGGCTCGGTG 82
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 584 CTGCTGACCCAGACGCGTTCGTATTCTTTCAGCCCGCGGCCACCCACAG 632
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 81 CTCCGGGAGGAGGAGCGCTCTCTCCGCTTCGCTGCGCGCGGTGGCCACG 33
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 33
B07644878/c
LOCUS
DEFINITION
    B07644878 566 bp mRNA linear EST 14-NOV-2002
    vegetative stage leaves Hordeum vulgare subsp. vulgare cDNA clone
    baak39j08 5', mRNA sequence.
ACCESSION
    B07644878
KEYWORDS
    EST.
SOURCE
    B07644878.1 GI:24963026
ORGANISM
    Hordeum vulgare subsp. vulgare
    Hordeum vulgare subsp. vulgare
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Poideae; Triticeae; Hordeum.
REFERENCE
    1 (bases 1 to 566)
AUTHORS
    Sato,K., Saisho,D. and Takeda,K.
TITLE
    Barley EST sequencing project in NIG and Okayama Univ
JOURNAL
    Unpublished (2002)
COMMENT
    Contact: Tadau Shin-i
    Center For Genetic Resource Information
    National Institute of Genetics
    1111 Yata, Mishima, Shizuoka 411-8540, Japan
    Tel: 81-559-81-6856
    Fax: 81-559-81-6855
    Email: tshini@genes.nig.ac.jp.
FEATURES
    source
    1..566
        /organism="Hordeum vulgare subsp. vulgare"
        /mol_type="mRNA"
        /cultivar="Akashinriki"
        /sub_species="vulgare"
        /db_xref="taxon:112509"
        /clone="baak39j08"
        /tissue_type="leaves"
        /dev_stage="vegetative stage"
        /clone_lib="K. Sato unpublished cDNA library, cv.
        Akashinriki vegetative stage leaves"
ORIGIN

Query Match          3.8%; Score 44.4; DB 4; Length 566;
Best Local Similarity 57.0%; Pred. No. 0.031;
Matches 81; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 386 CTTCTGAATGACATCAGGGATCCCGCGCTCTCACTGGCGATAACGGCAGCGCGGAGAC 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 CTCAGGAATAATGTCTAGGTATGCTTCAGCTGAGCAGCGACCCCGGAATCTCCGATGA 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 446 TGACGCTTCAGCGAGTACCATACCAAAACGCTTCATTTCCGAAGGATGACCAACCACT 505
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 CATTGCTCTCAACACACAAACCCAGCGCTCTGTGATTCAGAAGGCATCACAACACATC 446
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 506 GGCATATCCGGTAGCCGGTAAC 527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 445 CCCACTGGCATAGGCGCTGTGAC 424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 34
CK153409
LOCUS
DEFINITION
    CK153409 887 bp mRNA linear EST 05-DEC-2003
    FGAS032014 Triticum aestivum FGAS: Talt2 Triticum aestivum cDNA,
    mRNA sequence.
ACCESSION
    CK153409
KEYWORDS
    EST.
SOURCE
    CK153409.1 GI:38973446
    Triticum aestivum (bread wheat)
```

ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 887)
AUTHORS
Allard, P., Crosby, W.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
Penniket, C., Roach, J.L. and Sathan, F.
Functional Genomics of Abiotic Stress In Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033
Email: fgas_este@cs.u Sask.ca
This sequence is the direct result of the Base calling software
Phred (default parameters). It is the raw base calls. To aid in the
identification of the high quality insert the software Lucy
(default parameters) has been run on this sequence. Lucy identified
the region [123,646].
Plate: Talt239, row: L, column: 02.
FEATURES
Location/Qualifiers
source
1..887
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Wheat line CI 14106"
/db_xref="taxon:4565"
/lab_host="DH5 alpha"
/clone_lib="Triticum aestivum FGAS: Talt2"
/notes="Organ: Crown; Vector: pGEM-T; SSH (suppression
subtractive hybridization) cDNA library from genotype
CI14106 cold hardened at 2 C for 1 day (24 H) (tester) and
subtracted against genotype Norstar cold hardened at 2 C
for 21 days and 49 days (equal amount of cDNA pooled
together before subtraction, driver). Nitro-pyrole
anchored oligo-dT priming and non-directional cloning."
ORIGIN
Query Match 3.7%; Score 43.4; DB 7; Length 887;
Best Local Similarity 54.8%; Pred. No. 0.074;
Matches 86; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 371 CGTCCCTGTGCACATCTTCTGAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 430
DB 631 CTTTCTCTCTGATCTCTCAGTATATATGTCAGGTATGCCCCCAGCTCGAGCAGCAC 690
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTTCATTTCCGAAG 490
DB 691 CGGCACCCCGGATGACATGCTCTCAACACAGAAACAGCGTCTCTGACTCAGAAGG 750
QY 491 CATGACCCCACTGGCATCCGGTAGACCGGTAA 527
DB 751 CATCAAAACACATCTCCCTCGCATAGGCGCTGTGAC 787
RESULT 35
CG058902/c
LOCUS
DEFINITION
CG058902 281 bp DNA linear GSS 19-AUG-2003
PUIBY81TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTA0548N17,
genomic survey sequence.
ACCESSION
CG058902
VERSION
CG058902.1 GI:33931082
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 281)

AUTHORS
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUIBY81TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
source
1..281
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBTA0548N17"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"
ORIGIN
Query Match 3.6%; Score 42.4; DB 9; Length 281;
Best Local Similarity 54.5%; Pred. No. 0.11;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;
QY 371 CTTCCCTGTGCACATCTTCTGAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 430
DB 219 CTTCCCTCTCTGATCTCTCAGGTATATATGATATCTCTCGGCGAGCCCGACTAC 160
QY 431 GGGCAGCCGGAGACTGACGCTTCAGCCAGTACCATACCAACGCTTTCATTTCCGAAGG 490
DB 159 CGGAATCTCTGATGATGCTCTCCAGCACCACCAACCCAGTGCTCTGACTCGGAAGG 100
QY 491 CATGACCCCACTGGCATCCGGTAGACCGGTAA 526
DB 99 CATCAAAACACATCTCCCTCGCATAGGCGCTGTGA 64
RESULT 36
CA300278/c
LOCUS
DEFINITION
CA300278 431 bp mRNA linear EST 26-SEP-2003
SCRFLV1036H07.g LV1 Saccharum officinarum cDNA clone SCRFLV1036H07
5', mRNA sequence.
ACCESSION
CA300278
VERSION
CA300278.1 GI:36074129
KEYWORDS
EST.
SOURCE
Saccharum officinarum
ORGANISM
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
REFERENCE
1 (bases 1 to 431)
AUTHORS
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
CONTACT: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 036 row: H column: 07
Seq primer: T7 Promoter Primer.

cDNA library generated from (Stalk Bark from adult plants). cDNA was prepared from poly(A) mRNA using SuperScript Plasmid System kit (Invitrogen) and double-strand cDNAs were fractionated in 1% agarose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.uicamc.br/public/>

ORIGIN

	Query Match	3.6%	Score 42.4;	DB 6;	Length 599;
	Best Local Similarity	54.5%;	Pred. No. 0.14;		
	Matches	85;	Conservative 0;	Mismatches 71;	Indels 0; Gaps 0;
QY	371	CGTCCCTGTGCACATCTTCTGATGATCATCAGGGATCCGCGCGTCTCTCACTGGCGATAAC	430		
Db	200	CTTCCCTCTCTGATCCTCAGGTATGATCAGGTATTCCTCCAGCGGAGCCCCCATCTAC	141		
QY	431	GGGCAGCGCCGGAGACTGACGCTTTCAGCCAGTACCATACCAAAAGCTTTCATTTTCCGAAGG	490		
Db	140	CGGAATCTCCGATGACATGGCGCTCCAGCACAAACCAAAACCCAGCGTCTCTGACTCGGAAGG	81		
QY	491	CATGACCAACACACTGGGCAATCCGGTAGACCGGTAA	526		
Db	80	CATCACAAACACATATCCCGCTGGCATAGGCGCTGTGA	45		

RESULT 39	CA080408/c				
LOCUS		CA080408	616 bp	mRNA	linear
DEFINITION		SCVPAM1056E09.g AM1 Saccharum officinarum cDNA clone SCVPAM1056E09			EST 23-SEP-2003
					5', mRNA sequence.

CA080408	GI:3493327
CA080408.1	EST.
Saccharum officinarum	Saccharum officinarum
Saccharum officinarum	Saccharum officinarum
Embryophyta; Embryophyta; Tracheophyta; Eukaryota; Viridiplantae; Streptophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.	

[illegible]

```

1. .616
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCVPAM1056E09"
/lab_host="DH10B"
/clone_lib="AM1"
/note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSPot1; Site.1: SalI; Site.2:
NotI; An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from poly(A) mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-column and fragments sizing between 0.8 and

```

1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://succest.lad.ic.unicamp.br/public>

ORIGIN

	Query Match	3.6%	Score 42.4;	DB 6;	Length 615;
	Best Local Similarity	54.5%;	Pred. No. 0.14;		
	Matches	85;	Conservative	0;	Mismatches 71; Indels 0; Gaps 0;
Qy	371	CGTCCCTGTGCACATCTTCTGAATGATCATCAGGGATCCCGCCGTCCTCACTGGCGATAAC	430		
Db	187	CTTCCCTCTCTGATCTCTCAGTATGATATCAGGTATTCCTCCAGAGGAGCCCCCACTAC	128		
Qy	431	GGGCAAGCCCGGAGACTGACGCTTCAGCCAGTACCATACCAAGCGTTTCATTTCCGAAGG	490		
Db	127	CGGNACTCCCGATGACATGSCCTCCAGCAACAACAAACCAGCGTCTCTGACTCGGAAGG	68		
Qy	491	CATGACCAACACACTGGCAATCCGGTAGACCGGTAA	526		
Db	67	CATCAAAACAACATCCCGCTGSCATAGGCGCTGTGA	32		

RESULT 40
CA188291/c

CA188291	631 bp	linear	EST 24-SEP-2003
SCCCAM1C05C12.g	AM1	Saccharum officinarum	CDNA clone SCCCAM1C05C12
5', mRNA sequence.			

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.

FEATURES

```

1. .631
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCCCAM1C05C12"
/lab_host="DH10B"
/clone_lib="AM1"
/notes="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSport1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Detail
of each source of RNA and library construction can be
obtained at http://sucfest.lad.ic.unicamp.br/public"

```

ORIGIN

Query Match 3.6%; Score 42.4; DB 6; Length 631;
Best Local Similarity 54.5%; Pred. No. 0.14;
Matches 85; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY	371	CGTCCCTGTACACATTCTTCTGAATGACATCAGGGATCCCCCGTCTCACTGGCGATAAC	430
Db	187	CTTCCCTCTCTGATCCTCAGGTATGATATCAGGTATTCTCCAGAGCGAGCCCCCACTAC	128
QY	431	GGGCAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAACGCTTTCATTTTCCGAAGG	490
Db	127	CGGAATCTCCGATGACATGGCCCTCCAGCAACAARACCCAGCGTCTCTGACTCGGAAGG	68
QY	491	CATGACCAACACACTGGCAATCCGGTAGACCGGTAA	526
Db	67	CATCAAAACACATCCCCCGCTGGCATAGGCTGTGA	32

Search completed: June 4, 2005, 15:16:11
Job time : 4427 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 14:02:31 ; Search time 5404 Seconds
(without alignments)
10589.491 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctgacagagatggaagaaaa.....ttttactttttctctgcag 1181

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1181	100.0	1181	6	BD205212 Nucleotid
2	1181	100.0	1181	6	AX011298 Sequence
3	1018	86.2	7395	1	EC74P0157
4	1018	86.2	7395	6	Y11275 E.coli 7.4
5	1018	86.2	92077	1	AX191728 Sequence
6	1018	86.2	92077	6	AF074613 Escherich
7	1018	86.2	92077	6	AX191727 Sequence
8	1018	86.2	92721	1	AB011549 Escherich
9	916	77.6	5612	1	AF043470 Escherich
10	31	2.6	31	6	BD205235 Nucleotid
11	31	2.6	31	6	BD205236 Nucleotid
12	31	2.6	31	6	AX011321 Sequence
13	31	2.6	31	6	AX011322 Sequence
14	24	2.0	342905	1	EX571861 Phototrab
15	24	2.0	349980	6	AX770908 Sequence
16	23	1.9	301761	1	AE016945 Bacteroid
17	22	1.9	22	6	BD205231 Nucleotid
18	22	1.9	22	6	BD205232 Nucleotid
19	22	1.9	22	6	BD205234 Nucleotid

20	22	1.9	22	6	AX011317	AX011317 Sequence
21	22	1.9	22	6	AX011318	AX011318 Sequence
22	22	1.9	22	6	AX011320	AX011320 Sequence
23	22	1.9	1943	9	HSMB00322	AL049975 Homo sapi
24	22	1.9	134649	9	AC004047	AC004047 Homo sapi
25	22	1.9	147849	9	AC139366	AC139366 Homo sapi
26	22	1.9	176467	9	AC074349	AC074349 Homo sapi
27	22	1.9	189326	9	AC018462	AC018462 Homo sapi
28	22	1.9	189786	10	AC140370	AC140370 Mus muscu
29	22	1.9	213581	10	AC147034	AC147034 Mus muscu
30	22	1.9	235284	2	AC134760	AC134760 Rattus no
31	22	1.9	236365	2	AC087099	AC087099 Mus muscu
32	22	1.9	262998	2	AC109723	AC109723 Rattus no
33	22	1.9	293873	2	AC114694	AC114694 Rattus no
34	21	1.8	21	6	BD205237	BD205237 Nucleotid
35	21	1.8	21	6	AX011323	AX011323 Sequence
36	21	1.8	1337	5	BX930985	BX930985 Gallus ga
37	21	1.8	110000	1	AE017225	Continuation (41 o
38	21	1.8	110000	1	AE017334	Continuation (41 o
39	21	1.8	134257	5	BX005321	BX005321 Zebrafish
40	21	1.8	141721	2	AC149953	AC149953 Strongylo
41	21	1.8	162900	2	AC144798	AC144798 Mus muscu
42	21	1.8	178366	2	AC142217	AC142217 Rattus no
43	21	1.8	179399	5	AL929338	AL929338 Zebrafish
44	21	1.8	203595	2	CR385054	CR385054 Danio rer
45	21	1.8	210380	10	AL627238	AL627238 Mouse DNA
46	21	1.8	220613	2	AC097940	AC097940 Rattus no
47	21	1.8	235948	2	AC099574	AC099574 Mus muscu
48	21	1.8	262665	2	AC095462	AC095462 Rattus no
49	21	1.8	278299	2	AC129681	AC129681 Rattus no
50	21	1.8	290206	1	AE017037	AE017037 Bacillus
51	20	1.7	20	6	BD205233	BD205233 Nucleotid
52	20	1.7	20	6	AX011319	AX011319 Sequence
53	20	1.7	636	8	CNS01B66	AL113830 Botrytis
54	20	1.7	696	6	CNS01CUI	AL116002 Botrytis
55	20	1.7	1077	6	AX416175	AX416175 Sequence
56	20	1.7	3589	1	SCSECA	X79725 S. carnosus
57	20	1.7	3589	6	BD195978	BD195978 Antisense
58	20	1.7	3589	6	AR382769	AR382769 Sequence
59	20	1.7	5402	6	CQ590837	CQ590837 Sequence
60	20	1.7	18163	10	BX545919	BX545919 Mouse DNA
61	20	1.7	53339	9	AC091847	AC091847 Homo sapi
62	20	1.7	66540	9	AC008624	AC008624 Homo sapi
63	20	1.7	76072	10	BX119961	BX119961 Mouse DNA
64	20	1.7	79381	5	BX936290	BX936290 Zebrafish
65	20	1.7	83596	2	AC014135	AC014135 Drosophil
66	20	1.7	110000	1	AE017261	Continuation (7 of
67	20	1.7	110000	1	AE017261	Continuation (8 of
68	20	1.7	118985	9	AC094102	AC094102 Homo sapi
69	20	1.7	119730	9	AC117472	AC117472 Homo sapi
70	20	1.7	125174	9	AC103372	AC103372 Homo sapi
71	20	1.7	127927	2	AC122181	AC122181 Mus muscu
72	20	1.7	128449	2	AC148508	AC148508 Carollia
73	20	1.7	139628	5	AL929509	AL929509 Zebrafish
74	20	1.7	140904	2	AC119837	AC119837 Mus muscu
75	20	1.7	150934	2	AC022252	AC022252 Homo sapi
76	20	1.7	152086	2	AC121444	AC121444 Rattus no
77	20	1.7	157023	9	AL365360	AL365360 Human DNA
78	20	1.7	157457	3	AC092493	AC092493 Drosophil
79	20	1.7	161345	2	AC091968	AC091968 Homo sapi
80	20	1.7	162748	2	AC140263	AC140263 Mus muscu
81	20	1.7	162748	2	AC140263	AC140263 Mus muscu
82	20	1.7	164755	2	BX927094	BX927094 Danio rer
83	20	1.7	168615	3	AC022347	AC022347 Drosophil
84	20	1.7	168672	2	AC118984	AC118984 Papio anu
85	20	1.7	173450	9	AC091769	AC091769 Homo sapi
86	20	1.7	173535	2	AC120202	AC120202 Papio anu
87	20	1.7	175449	2	BX927146	BX927146 Danio rer
88	20	1.7	177472	9	AC008786	AC008786 Homo sapi
89	20	1.7	188485	2	AC027454	AC027454 Homo sapi
90	20	1.7	196966	10	AC087898	AC087898 Mus muscu
91	20	1.7	198189	2	AC145043	AC145043 Papio anu
92	20	1.7	200131	9	AC105213	AC105213 Homo sapi

93	20	1.7	201923	10	AC121772	AC121772 Mus muscu	c 166	19	1.6	131188	9	HSJ1103B4	AL121998 Human DNA
94	20	1.7	202704	10	AC090843	AC090843 Mus muscu	c 167	19	1.6	133195	8	AC007123	AC007123 Arabidops
95	20	1.7	211163	2	AC141508	AC141508 Rattus no	168	19	1.6	134979	9	AC018720	AC018720 Homo sapi
96	20	1.7	215362	10	AL663108	AL663108 Mouse DNA	169	19	1.6	136448	2	AC133662	AC133662 Rattus no
97	20	1.7	219564	2	AC073708	AC073708 Mus muscu	170	19	1.6	141288	2	AC068878	AC068878 Homo sapi
98	20	1.7	222150	2	AC102425	AC102425 Mus muscu	171	19	1.6	142344	2	AC128740	AC128740 Rattus no
99	20	1.7	226909	2	AC079490	AC079490 Mus muscu	c 172	19	1.6	143165	2	AL591114	AL591114 Homo sapi
100	20	1.7	227719	2	AC102776	AC102776 Mus muscu	173	19	1.6	143453	9	HSJ661G12	AL049797 Human DNA
101	20	1.7	228315	2	AC094380	AC094380 Rattus no	c 174	19	1.6	143864	2	AC048359	AC048359 Homo sapi
102	20	1.7	239439	10	AC117573	AC117573 Mus muscu	c 175	19	1.6	143981	9	HSJ1050K3	AL121776 Human DNA
103	20	1.7	241464	2	AC125803	AC125803 Rattus no	176	19	1.6	144092	5	AC092081	AC092081 Gallus ga
104	20	1.7	244845	2	AC105622	AC105622 Rattus no	177	19	1.6	144369	5	AC092403	AC092403 Gallus ga
105	20	1.7	248158	2	AC111483	AC111483 Rattus no	178	19	1.6	144386	2	AC120515	AC120515 Bos tauru
106	20	1.7	272065	2	AC117029	AC117029 Rattus no	c 179	19	1.6	145618	9	AC005138	AC005138 Homo sapi
107	20	1.7	290507	1	AE017327	AE017327 Listeria	c 180	19	1.6	145918	9	AC034229	AC034229 Homo sapi
108	20	1.7	313450	1	AL596170	AL596170 Listeria	c 181	19	1.6	146301	9	AC114494	AC114494 Homo sapi
109	20	1.7	314205	3	AE003505	AE003505 Drosophil	c 182	19	1.6	146950	9	AC093747	AC093747 Homo sapi
110	20	1.7	319630	6	AX413016	AX413016 Sequence	c 183	19	1.6	147844	2	AC068188	AC068188 Homo sapi
111	20	1.7	347050	1	AL591981	AL591981 Listeria	c 184	19	1.6	152209	2	AL591023	AL591023 Homo sapi
112	20	1.7	349980	6	AX417045	AX417045 Sequence	185	19	1.6	152416	9	AC096663	AC096663 Homo sapi
113	20	1.7	349980	6	AX417046	AX417046 Sequence	186	19	1.6	153307	10	AL845441	AL845441 Mouse DNA
114	20	1.7	349980	6	AX641669	AX641669 Sequence	187	19	1.6	153441	9	BS000071	BS000071 Pan trogl
115	19	1.6	199	9	HS76A10F	Z62955 H.sapiens C	188	19	1.6	153616	10	AC125105	AC125105 Mus muscu
116	19	1.6	201	11	BV170175	BV170175 sqm30358	c 189	19	1.6	155006	2	AC141681	AC141681 Apis mell
117	19	1.6	262	6	AX310258	AX310258 Sequence	c 190	19	1.6	156002	2	AC036110	AC036110 Homo sapi
118	19	1.6	267	6	AR557281	AR557281 Sequence	191	19	1.6	156413	2	AL365177	AL365177 Homo sapi
119	19	1.6	731	11	BV077066	BV077066 S212P6047	192	19	1.6	156506	2	AC068038	AC068038 Homo sapi
120	19	1.6	940	3	DME426910	AJ426910 Drosophil	c 193	19	1.6	157113	2	AC009054	AC009054 Homo sapi
121	19	1.6	1148	9	BC082970	BC082970 Homo sapi	c 194	19	1.6	157145	2	AC146323	AC146323 Pan trogl
122	19	1.6	1642	1	SHFVIRK	DL1025 Shigella fl	c 195	19	1.6	157248	8	AP004300	AP004300 Oryza sat
123	19	1.6	1642	6	R06635	E06635 Shigella vi	c 196	19	1.6	157493	2	AC027068	AC027068 Homo sapi
124	19	1.6	2680	6	CO727621	CO727621 Sequence	197	19	1.6	158216	2	AC024463	AC024463 Homo sapi
125	19	1.6	2806	9	HSA011863	AJ011863 Homo sapi	c 198	19	1.6	159887	8	OSJN00095	AL606652 Oryza sat
126	19	1.6	2982	6	AR338816	AR338816 Sequence	c 199	19	1.6	160380	9	AC124066	AC124066 Homo sapi
127	19	1.6	3815	1	AY206446	AY206446 Shigella	c 200	19	1.6	161194	2	AC021466	AC021466 Homo sapi
128	19	1.6	4314	6	Q0849732	Q0849732 Sequence	c 201	19	1.6	161326	9	AC004067	AC004067 Homo sapi
129	19	1.6	4314	6	Q0849732	Q0849732 Homo sapi	202	19	1.6	162877	9	AC018509	AC018509 Homo sapi
130	19	1.6	4523	5	XLA132556	AJ132556 Xenopus l	203	19	1.6	162995	10	AL663032	AL663032 Mouse DNA
131	19	1.6	11556	1	AE001713	AE001713 Thermotog	204	19	1.6	163621	9	AL445704	AL445704 Human DNA
132	19	1.6	13138	10	AF149203	AF149203 Mus muscu	205	19	1.6	163880	9	AC009505	AC009505 Homo sapi
133	19	1.6	15102	3	AY456187	AY456187 Argulus a	c 206	19	1.6	165378	2	CR381620	CR381620 Danio rer
134	19	1.6	36731	2	AC101042	AC101042 Mus muscu	c 207	19	1.6	165953	2	AC129243	AC129243 Rattus no
135	19	1.6	36963	9	AC144506	AC144506 Homo sapi	c 208	19	1.6	166787	9	AL353789	AL353789 Human DNA
136	19	1.6	37411	9	AC139077	AC139077 Homo sapi	c 209	19	1.6	167004	5	AL954840	AL954840 Zebrafish
137	19	1.6	39675	9	AP001117	AP001117 Homo sapi	210	19	1.6	167114	5	BX936464	BX936464 Zebrafish
138	19	1.6	42307	9	AC010510	AC010510 Homo sapi	211	19	1.6	167475	10	AC090881	AC090881 Mus Muscu
139	19	1.6	47217	9	AC108934	AC108934 Homo sapi	212	19	1.6	168065	9	CNS050TC6	AL355072 Human chr
140	19	1.6	69271	2	AC109313	AC109313 Homo sapi	c 213	19	1.6	168872	2	AC073325	AC073325 Homo sapi
141	19	1.6	69520	2	AC123272	Continuation (4 of	214	19	1.6	168919	9	AC101972	AC101972 Mus muscu
142	19	1.6	74634	2	AC09727	Continuation (4 of	c 215	19	1.6	169494	9	AC007630	AC007630 Homo sapi
143	19	1.6	76287	9	AL645617	AL645617 Human DNA	216	19	1.6	169739	10	AC112949	AC112949 Mus muscu
144	19	1.6	76881	9	AL645734	AL645734 Human DNA	c 217	19	1.6	170257	2	AC119604	AC119604 Rattus no
145	19	1.6	87788	9	AC087888	AC087888 Homo sapi	218	19	1.6	170529	9	AC008572	AC008572 Homo sapi
146	19	1.6	89377	8	AP006111	AP006111 Lotus cor	219	19	1.6	170601	9	AC122151	AC122151 Macaca mu
147	19	1.6	95419	2	OSIG00051	AL732356 Oryza sat	c 220	19	1.6	170781	2	AC092212	AC092212 Trypanoso
148	19	1.6	96444	9	HSJ738116	AL049867 Human DNA	221	19	1.6	171051	2	AC023528	AC023528 Homo sapi
149	19	1.6	102927	9	AC123964	AC123964 Macaca mu	c 222	19	1.6	171329	2	AC147785	AC147785 Pan trogl
150	19	1.6	103234	2	AP003997	AP003997 Oryza sat	c 223	19	1.6	171965	2	AC027507	AC027507 Homo sapi
151	19	1.6	107931	8	AC007478	AC007478 Arabidops	c 224	19	1.6	172105	9	AC093135	AC093135 Pan trogl
152	19	1.6	110000	2	AC091339	Continuation (10 of	c 225	19	1.6	172982	5	BX323052	BX323052 Zebrafish
153	19	1.6	110000	2	AC115674	Continuation (3 of	226	19	1.6	174327	9	AC007091	AC007091 Homo sapi
154	19	1.6	110000	2	AC129007	AC129007 Rattus no	227	19	1.6	174344	2	AC146763	AC146763 Zea mays
155	19	1.6	110000	2	AC151843	Continuation (3 of	c 228	19	1.6	175226	2	AC141861	AC141861 Gallus ga
156	19	1.6	110000	2	AF546187	Continuation (3 of	229	19	1.6	178625	9	AC106017	AC106017 Homo sapi
157	19	1.6	110401	9	AC004068	AC004068 Homo sapi	c 230	19	1.6	178847	10	AC115708	AC115708 Mus muscu
158	19	1.6	112031	9	AL445487	AL445487 Human DNA	231	19	1.6	180015	10	AC121925	AC121925 Mus muscu
159	19	1.6	117737	8	F6D8	AC008016 Arabidops	c 232	19	1.6	180104	9	DJ526N18	AF123462 Homo sapi
160	19	1.6	119008	9	AC013244	AC013244 Homo sapi	c 233	19	1.6	180399	2	AC079252	AC079252 Homo sapi
161	19	1.6	119420	8	HS97K18	AL021406 Human DNA	c 234	19	1.6	180673	10	AC068627	AC068627 Mus Muscu
162	19	1.6	123508	8	AC337620	AC337620 Oryza sat	c 235	19	1.6	182509	10	AC115706	AC115706 Mus muscu
163	19	1.6	123766	8	CNS08C8E	AL731874 Oryza sat	c 236	19	1.6	182586	2	AC053505	AC053505 Homo sapi
164	19	1.6	125653	9	AL133517	AL133517 Human DNA	c 237	19	1.6	182807	2	AC026544	AC026544 Homo sapi
165	19	1.6	127666	2	AC105069	AC105069 Mus muscu	238	19	1.6	183092	5	AC145912	AC145912 Gallus ga

239	19	1.6	183105	9	AC009497	Homo sapi	c 312	19	1.6	239870	2	AC094970	AC094970 Rattus no
240	19	1.6	183376	2	AC129621	Rattus no	c 313	19	1.6	239946	2	AC103505	AC103505 Rattus no
c 241	19	1.6	183915	2	AC073967	Homo sapi	314	19	1.6	240507	2	AC103202	AC103202 Rattus no
c 242	19	1.6	184463	10	AC126553	Mus muscu	315	19	1.6	242745	10	AC114995	Mus muscu
243	19	1.6	184925	2	AC027608	Homo sapi	c 316	19	1.6	243217	2	AC131403	Rattus no
c 244	19	1.6	186535	9	AC022272	Homo sapi	c 317	19	1.6	243777	2	AC131226	Rattus no
245	19	1.6	186670	9	AC080188	Homo sapi	c 318	19	1.6	247528	10	AL807756	Mouse DNA
246	19	1.6	187117	9	AC016749	Homo sapi	319	19	1.6	248320	2	AC134731	Rattus no
c 247	19	1.6	187117	9	AC016749	Homo sapi	320	19	1.6	248464	2	AC096438	Rattus no
c 248	19	1.6	187383	9	AC138954	Homo sapi	321	19	1.6	251242	2	AC110710	Rattus no
c 249	19	1.6	188165	9	CNS000003	Human chr	c 322	19	1.6	256124	2	AC130447	Rattus no
250	19	1.6	188345	9	AC092803	Homo sapi	c 323	19	1.6	256754	2	AC109139	Mus muscu
251	19	1.6	188630	2	AC012605	Homo sapi	c 324	19	1.6	256769	2	AC106589	Rattus no
252	19	1.6	188724	10	AC142113	Mus muscu	c 325	19	1.6	258073	2	AC094029	Rattus no
c 253	19	1.6	188977	2	AC145995	Pan trogl	326	19	1.6	258099	2	AC128361	Rattus no
c 254	19	1.6	190270	2	EX927285	Danio rer	327	19	1.6	258149	2	AC098277	Rattus no
255	19	1.6	191206	10	AC124573	Mus muscu	c 328	19	1.6	258649	2	AC112297	Rattus no
256	19	1.6	191906	10	AC132304	Mus muscu	c 329	19	1.6	258938	2	AC099416	Mus muscu
257	19	1.6	191918	2	AC026484	Homo sapi	c 330	19	1.6	260752	2	AC119315	Rattus no
258	19	1.6	192001	2	AC086836	Homo sapi	c 331	19	1.6	261309	2	AC099244	Rattus no
c 259	19	1.6	192349	2	AC115058	Mus muscu	c 332	19	1.6	265125	2	AC126716	Rattus no
c 260	19	1.6	192943	9	AC012645	Homo sapi	333	19	1.6	265136	2	AC094847	Rattus no
c 261	19	1.6	193460	2	AC112053	Rattus no	334	19	1.6	269654	2	AC098068	Rattus no
c 262	19	1.6	193665	9	AC105314	Homo sapi	335	19	1.6	274839	2	AC120723	Rattus no
263	19	1.6	193826	9	AL513479	Human DNA	336	19	1.6	278206	2	AL513471	Homo sapi
c 264	19	1.6	198433	10	AC127327	Mus muscu	337	19	1.6	281021	2	AC111415	Rattus no
c 265	19	1.6	198838	2	AC093647	Homo sapi	c 338	19	1.6	281021	2	AC111415	Rattus no
266	19	1.6	199103	2	AC080076	Homo sapi	c 339	19	1.6	283615	2	AC104295	Mus muscu
267	19	1.6	199271	2	AC142243	Mus muscu	c 340	19	1.6	288358	2	AC106121	Rattus no
268	19	1.6	199316	9	AC011373	Homo sapi	341	19	1.6	300201	2	AC107086	Rattus no
c 269	19	1.6	199670	9	AC142303	Pan trogl	342	19	1.6	300450	1	AP005371	Thermosyn
c 270	19	1.6	201328	10	AL840635	Mouse DNA	343	19	1.6	302212	1	AB017147	Helicobac
c 271	19	1.6	201399	2	AC123392	Rattus no	c 344	19	1.6	309838	2	AC118433	Rattus no
272	19	1.6	201733	2	CR391953	Danio rer	c 345	19	1.6	313113	2	AC113623	Rattus no
273	19	1.6	202455	2	AC127814	Rattus no	c 346	19	1.6	321034	2	AC118305	Rattus no
274	19	1.6	203560	2	AC026298	Homo sapi	c 347	19	1.6	330580	2	AC101959	Mus muscu
c 275	19	1.6	203560	2	AC026298	Homo sapi	348	19	1.6	340000	9	HS21C027	Homo sapi
c 276	19	1.6	208185	2	AC009105	Homo sapi	c 349	19	1.6	346792	1	BX571658	Wolfinella
c 277	19	1.6	208211	10	AC108797	Mus muscu	c 350	19	1.6	348885	2	AC151274	Mus muscu
278	19	1.6	208457	2	AC141960	Rattus no	351	18	1.5	160	3	AY654016	Panaeus m
279	19	1.6	209896	5	BX248137	Zebrafish	352	18	1.5	221	10	AF200939	Meriones
c 280	19	1.6	210450	2	AC128567	Rattus no	353	18	1.5	240	8	AY023952	Oryza sat
c 281	19	1.6	213452	10	AC077689	Mus muscu	c 354	18	1.5	453	6	CQ752408	Sequence
c 282	19	1.6	213462	1	SFPWR100	Shigella	c 355	18	1.5	480	9	HS323155	Homo sapi
c 283	19	1.6	214270	10	AC133908	Mus muscu	c 356	18	1.5	497	4	AB098915	Boa tauru
c 284	19	1.6	215314	10	AC140373	Mus muscu	357	18	1.5	499	4	AB098765	Boa tauru
c 285	19	1.6	215400	10	AC115937	Mus muscu	c 358	18	1.5	531	6	AR506660	Sequence
c 286	19	1.6	216882	2	AC132716	Rattus no	c 359	18	1.5	535	5	AY101222	Uraetoph
c 287	19	1.6	217001	10	AC103419	Rattus no	360	18	1.5	540	5	AY101223	Uraetoph
c 288	19	1.6	219462	2	AC151574	Mus muscu	361	18	1.5	580	4	AB098768	Boa tauru
289	19	1.6	219784	10	AC112792	Mus muscu	362	18	1.5	666	4	AB098831	Boa tauru
c 290	19	1.6	220226	2	AC110124	Rattus no	c 363	18	1.5	672	4	AF540563	Boa tauru
c 291	19	1.6	220976	10	AC132474	Mus muscu	c 364	18	1.5	675	11	BV050003	Ovis aries
c 292	19	1.6	221091	9	AC146228	Pan trogl	c 365	18	1.5	775	4	OAUS4800	Sequence
c 293	19	1.6	221232	2	AC148542	Mus muscu	c 366	18	1.5	824	8	YSCDCAA	S.cerevisia
c 294	19	1.6	221530	2	AC138095	Mus muscu	c 367	18	1.5	864	4	AB003093	Boa tauru
c 295	19	1.6	221618	1	AF386526	Shigella	c 368	18	1.5	954	6	AR395267	Sequence
296	19	1.6	221622	10	AC125183	Mus muscu	369	18	1.5	957	4	AF038313	Antechinu
c 297	19	1.6	221851	1	AF348706	Shigella	c 370	18	1.5	1031	1	ECDDPAGEN	E.cloacae d
c 298	19	1.6	222236	2	AC096440	Rattus no	c 371	18	1.5	1154	6	AR165145	Sequence
299	19	1.6	222895	2	AC084065	Mus muscu	c 372	18	1.5	1496	10	AF313617	Mus muscu
c 300	19	1.6	223879	9	AC008735	Homo sapi	c 373	18	1.5	1671	8	SCYLL196C	S.cerevisia
c 301	19	1.6	231967	10	AC122091	Rattus no	c 374	18	1.5	2154	8	AR396413	Sequence
c 302	19	1.6	232410	2	AC129999	Rattus no	c 375	18	1.5	2170	6	BD193949	Enterococ
303	19	1.6	233358	2	AC095317	Rattus no	376	18	1.5	2281	6	AR193027	Sequence
c 304	19	1.6	234053	2	AC128262	Rattus no	c 377	18	1.5	2344	5	BC067712	Danio rer
305	19	1.6	234508	2	AC113757	Rattus no	c 378	18	1.5	2443	8	SCDC6G	S.cerevisia
306	19	1.6	235682	2	AC126709	Rattus no	379	18	1.5	2583	8	SCYLL195C	Sequence
307	19	1.6	235938	2	AC121469	Rattus no	380	18	1.5	2630	8	YSCDC6	S.cerevisia
308	19	1.6	236528	2	AC132013	Rattus no	381	18	1.5	2894	6	AR338604	Sequence
309	19	1.6	237468	2	AC108532	Rattus no	382	18	1.5	2895	6	CQ720031	Sequence
310	19	1.6	238955	2	AC126169	Rattus no	383	18	1.5	2974	6	BD193828	Enterococ
c 311	19	1.6	238969	2	AC134114	Rattus no	384	18	1.5	3007	10	RNAJ1933	Rattus no

385	18	1.5	3037	10	AF110416	AF110416 Rattus no	458	18	1.5	56539	2	AC083758	AC083758 Homo sapi
386	18	1.5	3038	10	AB017260	AB017260 Rattus no	C 459	18	1.5	57074	8	AP068976	AP068976 Homo sapi
C 387	18	1.5	3530	4	BTAS18965	BTAS18965 Bos tauru	C 460	18	1.5	58137	9	AP004590	AP004590 Oryza sat
C 388	18	1.5	3532	6	CQ614303	CQ614303 Sequence	C 461	18	1.5	58289	5	BX470142	BX470142 Zebrafish
389	18	1.5	3686	6	AX834555	AX834555 Sequence	C 462	18	1.5	59580	2	AC106875	AC106875 Homo sapi
390	18	1.5	3686	9	AK097086	AK097086 Homo sapi	C 463	18	1.5	59620	2	AC014068	AC014068 Drosophill
391	18	1.5	3810	3	DMU01088	DMU01088 Drosophilla	C 464	18	1.5	60837	9	HSJ354112	AL078596 Human DNA
C 392	18	1.5	3971	9	BC051694	BC051694 Homo sapi	C 465	18	1.5	64183	9	HSJ429G5	AL078596 Human DNA
393	18	1.5	4083	3	BOL547623	ALJ547623 Bactrocerc	C 466	18	1.5	65190	2	AC036124	AL442109 Oryza sat
394	18	1.5	4720	3	DMU01087	DMU01087 Drosophilla	C 467	18	1.5	65942	2	AC036124	AL442109 Oryza sat
395	18	1.5	5024	10	BC067398	BC067398 Mus muscu	C 468	18	1.5	66324	8	NCB99J10	AL356324 Neurospor
396	18	1.5	5535	3	AB000802	AB000802 Caenorhab	C 469	18	1.5	69376	2	AC016022	AL356324 Neurospor
C 397	18	1.5	6803	1	EC0576011	EC0576011 Escherich	C 470	18	1.5	69520	10	AL669925	AL669925 Mouse DNA
C 398	18	1.5	8894	6	CQ872947	CQ872947 Sequence	C 471	18	1.5	70172	2	AC1130370	AL669925 Mouse DNA
C 399	18	1.5	9566	1	RP1W17A	AL0585 Xanthomonas	C 472	18	1.5	70299	9	AL161945	AL161945 Homo sapi
400	18	1.5	9652	1	AE015181	AE015181 Shigella	C 473	18	1.5	71437	2	AC026919	AL161945 Homo sapi
C 401	18	1.5	10029	1	AE011543	AE011543 Leptospir	C 474	18	1.5	71437	2	AC026919	AL161945 Homo sapi
402	18	1.5	10029	1	AE011543	AE011543 Streptoco	C 475	18	1.5	71437	2	AC026919	AL161945 Homo sapi
403	18	1.5	10227	1	U32725	U32725 Haemophilus	C 476	18	1.5	72231	2	AC124284	AC025810 Homo sapi
C 404	18	1.5	10382	1	AE007611	AE007611 Clostridi	C 477	18	1.5	72969	2	AC027724	AC124284 Homo sapi
C 405	18	1.5	10551	1	AE012326	AE012326 Xanthomon	C 478	18	1.5	73094	2	AL671114	AC027724 Homo sapi
C 406	18	1.5	10904	1	AY079086	AY079086 Corynebac	C 479	18	1.5	74037	9	HSJ071N3	AL671114 Homo sapi
407	18	1.5	10952	1	AE005382	AE005382 Escherich	C 480	18	1.5	76362	2	AC133451_3	AL031728 Human DNA
C 408	18	1.5	11573	1	AE010024	AE010024 Streptoco	C 481	18	1.5	78647	2	AC108346_3	Continuation (4 of
C 409	18	1.5	12230	1	AE006747	AE006747 Sulfolobu	C 482	18	1.5	78743	2	CR753815	Continuation (4 of
C 410	18	1.5	12394	14	AF195822	AF195822 Grapevine	C 483	18	1.5	80307	9	AC128684	CR753815 Homo sapi
411	18	1.5	12663	2	AC012809	AC012809 Drosophill	C 484	18	1.5	80547	9	AC007382	AC128684 Homo sapi
C 412	18	1.5	13377	8	AF541860S11	AF541870 Chlamydom	C 485	18	1.5	80818	8	AB019232	AC128684 Homo sapi
C 413	18	1.5	14798	9	AL162429	AL162429 Human DNA	C 486	18	1.5	81778	2	AC136102	AB019232 Arabidops
414	18	1.5	14951	1	AE000756	AE000756 Aquifex a	C 487	18	1.5	83521	9	AC129510	AC136102 Rattus no
415	18	1.5	15312	1	D90801	D90801 E.coli geno	C 488	18	1.5	84742	2	AC099937	AC129510 Homo sapi
416	18	1.5	16432	5	AY456254	AY456254 Uraeotlyph	C 489	18	1.5	86129	2	BX928745	AC099937 Mus muscu
417	18	1.5	17300	1	D90800	D90800 E.coli geno	C 490	18	1.5	87133	2	AC098640_3	BX928745 Danio rer
418	18	1.5	19291	1	D90802	D90802 E.coli geno	C 491	18	1.5	90019	8	TM021804_3	Continuation (4 of
C 419	18	1.5	20382	1	AE008856	AE008856 Salmonell	C 492	18	1.5	91840	9	AC112254	Continuation (4 of
C 420	18	1.5	20524	1	D90799	D90799 E.coli geno	C 493	18	1.5	92911	2	AC020218	AC112254 Homo sapi
C 421	18	1.5	20663	2	AC014972	AC014972 Drosophill	C 494	18	1.5	94341	9	AL580336	AC020218 Drosophill
422	18	1.5	21233	9	CR759801	CR759801 Human DNA	C 495	18	1.5	94665	3	AC099765	AL580336 Human DNA
C 423	18	1.5	21267	3	CEV70D2A	AL033513 Caenorhab	C 496	18	1.5	95272	3	AC005298	AC099765 Caenorhab
424	18	1.5	27955	3	CEC04H5	Z81462 Caenorhabdi	C 497	18	1.5	95579	9	AC023162	AC005298 Drosophill
C 425	18	1.5	28026	4	AF125462	AF125462 Caenorhab	C 498	18	1.5	96066	9	AC135056	AC023162 Homo sapi
426	18	1.5	30014	4	AY152832	AY152832 Felis cat	C 499	18	1.5	96191	8	AP004159	AC135056 Homo sapi
427	18	1.5	30827	3	CEW09H1	Z82081 Caenorhabdi	C 500	18	1.5	96798	9	AL356420	AP004159 Oryza sat
C 428	18	1.5	31694	9	AL353606	AL353606 Human DNA	C 501	18	1.5	96960	5	AC146894	AL356420 Human DNA
C 429	18	1.5	32656	2	AC021533	AC021533 Homo sapi	C 502	18	1.5	97817	9	AC063922	AC146894 X. tropic
C 430	18	1.5	33811	2	AC101185	AC101185 Mus muscu	C 503	18	1.5	98356	9	HSJ344J2	AC063922 Homo sapi
C 431	18	1.5	34107	3	CEFO2H6	Z82265 Caenorhabdi	C 504	18	1.5	98415	9	AC073421	AL449213 Homo sapi
C 432	18	1.5	34336	3	U39851	U39851 Caenorhabdi	C 505	18	1.5	99592	9	AF263284	AC073421 Homo sapi
C 433	18	1.5	34679	2	AL671034	AL671034 Homo sapi	C 506	18	1.5	99923	9	AC079409	AF263284 Homo sapi
434	18	1.5	35054	9	HS333B10	Z81450 Human DNA s	C 507	18	1.5	100000	9	AB020869	AC079409 Homo sapi
435	18	1.5	36017	8	SCRRP21	X77688 S.cerevisia	C 508	18	1.5	100000	9	AP000518	AB020869 Homo sapi
436	18	1.5	36418	1	AC004652	AC004652 Homo sapi	C 509	18	1.5	100183	9	BX088647	AP000518 Homo sapi
437	18	1.5	37881	9	AE058689	AE058689 Escherich	C 510	18	1.5	101531	10	AL732357	BX088647 Human DNA
C 438	18	1.5	38784	2	AC012882	AC012882 Drosophill	C 511	18	1.5	102086	9	AL512447	AL732357 Mouse DNA
C 439	18	1.5	39337	3	U39848	U39848 Caenorhabdi	C 512	18	1.5	103922	9	AC096546	AL512447 Human DNA
C 440	18	1.5	40059	6	CQ593474	CQ593474 Sequence	C 513	18	1.5	104011	9	AL590392	AC096546 Homo sapi
441	18	1.5	40181	3	CEB0457	AL141931 Homo sapi	C 514	18	1.5	104296	8	CNS080C8H	AL590392 Human DNA
C 442	18	1.5	40634	3	CEB0457	Z54306 Caenorhabdi	C 515	18	1.5	104648	9	HSJ096108	AL731877 Oryza sat
C 443	18	1.5	43125	3	AF147779	AF147779 Drosophill	C 516	18	1.5	104695	9	HSJ096108	AL731877 Oryza sat
C 444	18	1.5	43284	9	AC004201	AC004201 Homo sapi	C 517	18	1.5	104872	10	AC084322	AL121879 Human DNA
445	18	1.5	43906	9	AL731824	AL731824 Human DNA	C 518	18	1.5	105113	8	AC151520	AC084322 Mus muscu
C 446	18	1.5	43956	3	AF098989	AF098989 Caenorhab	C 519	18	1.5	108490	9	AC118480	AC151520 Medicago
447	18	1.5	46296	5	BX276107	BX276107 Zebrafish	C 520	18	1.5	108623	5	BX649472	AC118480 Homo sapi
C 448	18	1.5	48657	9	AL354755	AL354755 Human DNA	C 521	18	1.5	109279	9	HSJ394F12	BX649472 Zebrafish
449	18	1.5	48843	8	AP467900	AP467900 Prunus pe	C 522	18	1.5	110000	1	U00096_16	Z83823 Human DNA s
450	18	1.5	51859	9	HSJ1193N1	AL121915 Human DNA	C 523	18	1.5	110000	1	AE016827_09	Continuation (17 o
451	18	1.5	52173	9	HSJ6PDGEN	X55448 H.sapiens G	C 524	18	1.5	110000	1	CP000003_07	Continuation (10 o
452	18	1.5	52928	9	BX322644	BX322644 Human DNA	C 525	18	1.5	110000	2	AC007194_2	Continuation (8 of
C 453	18	1.5	54356	2	AC083926	AC083926 Homo sapi	C 526	18	1.5	110000	2	AC098773_2	Continuation (3 of
454	18	1.5	54492	9	AL590703	AL590703 Human DNA	C 527	18	1.5	110000	2	AC107170_1	Continuation (2 of
455	18	1.5	55216	6	AR306445	AR306445 Sequence	C 528	18	1.5	110000	2	AC116854_2	Continuation (3 of
C 456	18	1.5	55216	6	AR306445	AR306445 Sequence	C 529	18	1.5	110000	2	AC118331_1	Continuation (2 of
C 457	18	1.5	56031	9	AL157377	AL157377 Human DNA	C 530	18	1.5	110000	2	AL928849_0	AL928849 Mus muscu

531	18	1.5	110000	6	BD426631_04	Continuation (5 of	604	18	1.5	145068	2	CR847956	CR847956
532	18	1.5	110000	6	AR274513_04	Continuation (5 of	605	18	1.5	145264	9	AC107939	AC107939 Homo sapi
533	18	1.5	110000	6	AR541453_04	Continuation (5 of	606	18	1.5	145458	9	AC150007	AC150007 Bos tauru
534	18	1.5	110261	9	AP006205	AP006209 Homo sapi	607	18	1.5	145726	2	AC021807	AC021807 Homo sapi
535	18	1.5	114104	9	AC128712	AC128712 Homo sapi	608	18	1.5	145846	2	AC020298	AC020298 Drosophi
536	18	1.5	114594	9	AC110999	AC110999 Homo sapi	c 609	18	1.5	145909	9	AC016482	AC016482 Homo sapi
537	18	1.5	114809	2	AC083823	AC083823 Homo sapi	c 610	18	1.5	146437	9	AC022836	AC022836 Homo sapi
c 538	18	1.5	115231	9	HS785G19	AL035458 Human DNA	611	18	1.5	146524	2	AC019040	AC019040 Homo sapi
539	18	1.5	115279	9	AC112506	AC112506 Homo sapi	c 612	18	1.5	147324	9	AF277315	AF277315 Homo sapi
540	18	1.5	116169	2	AP002080	AP002080 Homo sapi	c 613	18	1.5	147327	10	AL808011	AL808011 Mouse DNA
541	18	1.5	117576	6	Q873159	Q873159 Sequence	c 614	18	1.5	147330	9	AC093514	AC093514 Homo sapi
c 542	18	1.5	117945	2	AC149226	AC149226 Danio rer	c 615	18	1.5	147458	9	AL357832	AL357832 Human DNA
543	18	1.5	118510	9	AL731543	AL731543 Human DNA	c 616	18	1.5	147507	2	AC148808	AC148808 Macropus
c 544	18	1.5	119350	9	HS128O3	Z98742 Human DNA s	617	18	1.5	147859	2	AC011079	AC011079 Homo sapi
545	18	1.5	119950	9	HS626B19	AL035087 Human DNA	c 618	18	1.5	148435	9	AC073973	AC073973 Homo sapi
546	18	1.5	120477	2	AL355995	AL355995 Homo sapi	619	18	1.5	148872	2	AC073168	AC073168 Homo sapi
547	18	1.5	120955	9	HUAC002310	AC002310 Human Chr	c 620	18	1.5	149146	2	AL583851	AL583851 Homo sapi
c 548	18	1.5	121689	9	AP000478	AP000478 Homo sapi	c 621	18	1.5	149370	9	AC012626	AC012626 Homo sapi
549	18	1.5	121889	5	BX255952	BX255952 Zebrafish	c 622	18	1.5	149400	8	AC121327	AC121327 Oryza sat
550	18	1.5	122216	8	AP003811	AP003811 Oryza sat	c 623	18	1.5	149433	9	AL135901	AL135901 Human DNA
551	18	1.5	123115	10	AL645688	AL645688 Mouse DNA	c 624	18	1.5	149920	9	AC093754	AC093754 Homo sapi
c 552	18	1.5	124525	10	AC134332	AC134332 Mus muscu	c 625	18	1.5	150300	8	AP003878	AP003878 Oryza sat
c 553	18	1.5	125150	2	AC005450	AC005450 Drosophi	626	18	1.5	150372	2	AC150103	AC150103 Gallus ga
554	18	1.5	127307	2	AC151398	AC151398 Felis cat	c 627	18	1.5	151517	2	AC031989	AC031989 Homo sapi
555	18	1.5	128082	2	AC140683	AC140683 Rattus no	c 628	18	1.5	151923	2	AC146490	AC146490 Macaca mu
c 556	18	1.5	128615	2	AL450329	AL450329 Homo sapi	c 629	18	1.5	151941	9	AC108131	AC108131 Homo sapi
557	18	1.5	128875	2	AC115829	AC115829 Mus muscu	c 630	18	1.5	152086	9	AL355796	AL355796 Human DNA
c 558	18	1.5	129109	9	AL390239	AL390239 Human DNA	631	18	1.5	152594	9	AL158196	AL158196 Human DNA
559	18	1.5	132033	9	AC099796	AC099796 Homo sapi	632	18	1.5	152813	9	AC023827	AC023827 Homo sapi
560	18	1.5	132575	10	AC117261	AC117261 Mus muscu	633	18	1.5	153368	8	AC136222	AC136222 Oryza sat
c 561	18	1.5	133069	2	AC026836	AC026836 Homo sapi	634	18	1.5	153562	9	AC009046	AC009046 Homo sapi
562	18	1.5	134341	9	AC099794	AC099794 Homo sapi	635	18	1.5	153736	9	AC079301	AC079301 Homo sapi
563	18	1.5	135925	5	BX927203	BX927203 Zebrafish	c 636	18	1.5	153787	2	AC023037	AC023037 Homo sapi
564	18	1.5	136748	2	AC151196	AC151196 Bos tauru	c 637	18	1.5	154130	9	HS657E11	HS657E11 Human DNA
565	18	1.5	136753	2	AC149968	AC149968 Strongylo	c 638	18	1.5	154405	9	AC004990	AC004990 Homo sapi
c 566	18	1.5	137165	2	AC011264	AC011264 Homo sapi	c 639	18	1.5	154865	9	AC099757	AC099757 Homo sapi
c 567	18	1.5	137217	2	HS211110	AL672302 Homo sapi	c 640	18	1.5	154867	9	AC017035	AC017035 Homo sapi
568	18	1.5	137384	10	AC121893	AC121893 Mus muscu	c 641	18	1.5	155146	2	AC025216	AC025216 Homo sapi
c 569	18	1.5	137385	2	AC113357	AC113357 Homo sapi	642	18	1.5	155168	3	DMER42117	DMER42117 Drosophi
570	18	1.5	137831	2	AC008921	AC008921 Homo sapi	c 643	18	1.5	155221	9	AC145688	AC145688 Pan trogl
571	18	1.5	137858	2	AC151397	AC151397 Felis cat	c 644	18	1.5	155318	9	HS791K14	HS791K14 Human DNA
c 572	18	1.5	138024	2	AC131381	AC131381 Strongylo	c 645	18	1.5	155321	2	AC130699	AC130699 Mus muscu
573	18	1.5	138710	10	AC023608	AC023608 Mus muscu	c 646	18	1.5	155487	2	AC141549	AC141549 Rattus no
574	18	1.5	138829	8	AP004638	AP004638 Psilotom	c 647	18	1.5	155490	2	CR749749	CR749749 Danio rer
c 575	18	1.5	138829	8	AP004638	AP004638 Psilotom	c 648	18	1.5	155523	2	AC025367	AC025367 Homo sapi
576	18	1.5	138869	5	BX322554	BX322554 Zebrafish	649	18	1.5	155618	2	CR405702	CR405702 Danio rer
c 577	18	1.5	139123	9	AL513527	AL513527 Human DNA	c 650	18	1.5	155718	10	AC134326	AC134326 Mus muscu
c 578	18	1.5	139438	9	AC009424	AC009424 Homo sapi	c 651	18	1.5	155877	2	AC073631	AC073631 Homo sapi
c 579	18	1.5	139495	2	CR381598	CR381598 Danio rer	c 652	18	1.5	156096	3	AC023748	AC023748 Drosophi
580	18	1.5	139544	9	HS0885M6	AL121926 Human DNA	c 653	18	1.5	156339	2	AC148567	AC148567 Canis fam
c 581	18	1.5	139722	9	AC093770	AC093770 Homo sapi	654	18	1.5	156655	9	HSN14	HSN14 Homo sapi
c 582	18	1.5	139774	2	AC069339	AC069339 Homo sapi	c 655	18	1.5	156784	10	AC120366	AC120366 Mus muscu
-583	18	1.5	140454	8	OSJN00183	AL662984 Oryza sat	c 656	18	1.5	156821	9	AC005691	AC005691 Homo sapi
584	18	1.5	140531	2	AC105341	AC105341 Homo sapi	c 657	18	1.5	157205	8	AP005098	AP005098 Oryza sat
c 585	18	1.5	140564	9	AC073975	AC073975 Homo sapi	c 658	18	1.5	157493	10	AC120085	AC120085 Rattus no
586	18	1.5	140833	2	CR847549	CR847549 Danio rer	c 659	18	1.5	157534	2	AC090100	AC090100 Homo sapi
c 587	18	1.5	140915	9	AC005587	AC005587 Homo sapi	660	18	1.5	157568	9	AP005138	AP005138 Homo sapi
c 588	18	1.5	140947	5	CR384085	CR384085 Zebrafish	661	18	1.5	157679	9	AP003470	AP003470 Homo sapi
589	18	1.5	141214	4	AY152830	AY152830 Felis cat	c 662	18	1.5	158060	2	AC024423	AC024423 Homo sapi
c 590	18	1.5	141288	2	AC068878	AC068878 Homo sapi	c 663	18	1.5	158116	2	AL357622	AL357622 Homo sapi
c 591	18	1.5	141473	2	AC102207	AC102207 Mus muscu	c 664	18	1.5	158121	2	AC026594	AC026594 Homo sapi
c 592	18	1.5	142080	8	AC079748	AC079748 Homo sapi	665	18	1.5	158242	9	AC146111	AC146111 Pan trogl
593	18	1.5	142670	8	CNS08C9U	AL772416 Oryza sat	666	18	1.5	158405	9	AC009753	AC009753 Homo sapi
c 594	18	1.5	142965	9	AC087528	AC087528 Homo sapi	c 667	18	1.5	158489	2	AC087803	AC087803 Homo sapi
c 595	18	1.5	143744	9	AC124283	AC124283 Homo sapi	c 668	18	1.5	158591	10	AC093022	AC093022 Mus Muscu
596	18	1.5	143905	2	AC149969	AC149969 Strongylo	c 669	18	1.5	158682	9	AP005363	AP005363 Homo sapi
597	18	1.5	144169	9	AL671859	AL671859 Human DNA	c 670	18	1.5	158813	9	AC080091	AC080091 Homo sapi
c 598	18	1.5	144463	10	AL671922	AL671922 Mouse DNA	671	18	1.5	158828	2	BX936381	BX936381 Danio rer
599	18	1.5	144576	9	AC141057	AC141057 Homo sapi	c 672	18	1.5	158958	9	AC007611	AC007611 Homo sapi
c 600	18	1.5	144577	9	AC023824	AC023824 Homo sapi	c 673	18	1.5	159020	2	AC027023	AC027023 Homo sapi
c 601	18	1.5	144683	9	AC096554	AC096554 Homo sapi	c 674	18	1.5	159082	10	AC124408	AC124408 Mus muscu
c 602	18	1.5	144767	8	CNS08C8O	AL731884 Oryza sat	c 675	18	1.5	159123	9	AC007739	AC007739 Homo sapi
603	18	1.5	145055	3	AC007453	AC007453 Drosophi	676	18	1.5	159178	9	AL665914	AL665914 Human DNA

677	18	1.5	159469	10	AC121920	AC121920 Mus muscu	750	18	1.5	167755	2	BX957242	BX957242 Danio rer
678	18	1.5	159570	9	AL136455	AL136455 Human DNA	C 751	18	1.5	168000	2	AC062030	AC062030 Homo sapi
679	18	1.5	159617	10	AL929024	AL929024 Mouse DNA	C 752	18	1.5	168023	2	AC137786	AC137786 Homo sapi
680	18	1.5	159625	9	AC114811	AC114811 Homo sapi	C 753	18	1.5	168105	9	AC093815	AC093815 Homo sapi
681	18	1.5	159643	9	AC104380	AC104380 Homo sapi	C 754	18	1.5	168172	2	AC010528	AC010528 Homo sapi
682	18	1.5	159692	9	AC016727	AC016727 Homo sapi	C 755	18	1.5	168247	2	AC073025	AC073025 Homo sapi
683	18	1.5	159832	2	AC104199	AC104199 Mus muscu	C 756	18	1.5	168308	2	AC084775	AC084775 Homo sapi
684	18	1.5	159832	2	AC143837	AC143837 Macaca mu	C 757	18	1.5	168423	10	AC131773	AC131773 Mus muscu
685	18	1.5	160025	2	AC018746	AC018746 Homo sapi	C 758	18	1.5	168435	2	AC148842	AC148842 Orolemur
686	18	1.5	160197	2	AC009814	AC009814 Homo sapi	C 759	18	1.5	168508	2	AC026236	AC026236 Homo sapi
687	18	1.5	160259	5	AL954386	AL954386 Zebrafish	C 760	18	1.5	168637	9	AC073137	AC073137 Homo sapi
688	18	1.5	160262	2	AC005242	AC005242 Homo sapi	C 761	18	1.5	169020	2	AC148160	AC148160 Zea mays
689	18	1.5	160276	2	CR376832	CR376832 Danio rer	C 762	18	1.5	169587	2	AC123355	AC123355 Rattus no
690	18	1.5	160656	2	AL592072	AL592072 Homo sapi	C 763	18	1.5	169630	5	BX957309	BX957309 Zebrafish
691	18	1.5	160666	2	AC068959	AC068959 Homo sapi	C 764	18	1.5	170212	9	AC058822	AC058822 Homo sapi
692	18	1.5	160714	2	AC150149	AC150149 Gallus ga	C 765	18	1.5	170580	2	AC119205	AC119205 Mus muscu
693	18	1.5	160831	2	AC025015	AC025015 Homo sapi	C 766	18	1.5	170718	3	AC104514	AC104514 Drosophil
694	18	1.5	161056	2	CR388093	CR388093 Danio rer	C 767	18	1.5	170767	5	BX936423	BX936423 Zebrafish
695	18	1.5	161266	2	AC148285	AC148285 Rhinolph	C 768	18	1.5	170832	2	CR376769	CR376769 Danio rer
696	18	1.5	161331	10	AL611950	AL611950 Mouse DNA	C 769	18	1.5	170884	10	AC137556	AC137556 Mus muscu
697	18	1.5	161478	9	AC025538	AC025538 Homo sapi	C 770	18	1.5	170949	2	AL669850	AL669850 Mus muscu
698	18	1.5	161602	9	AP001978	AP001978 Homo sapi	C 771	18	1.5	171051	9	AC022960	AC022960 Homo sapi
699	18	1.5	161949	10	AC122386	AC122386 Mus muscu	C 772	18	1.5	171058	5	BX323033	BX323033 Zebrafish
700	18	1.5	162069	2	AC108928	AC108928 Homo sapi	C 773	18	1.5	171167	9	AC012354	AC012354 Homo sapi
701	18	1.5	162180	9	AC124287	AC124287 Homo sapi	C 774	18	1.5	171529	2	AC023021	AC023021 Homo sapi
702	18	1.5	162223	9	AC145988	AC145988 Pan trogl	C 775	18	1.5	171551	8	OSJN00177	AL662976 Oryza sat
703	18	1.5	162378	2	BX927101	BX927101 Danio rer	C 776	18	1.5	171681	2	AC091541	AC091541 Canis fam
704	18	1.5	162383	9	AB023054	AB023054 Homo sapi	C 777	18	1.5	171715	2	AC132398	AC132398 Mus muscu
705	18	1.5	162393	9	AB023055	AB023055 Homo sapi	C 778	18	1.5	171852	9	AL137069	AL137069 Human DNA
706	18	1.5	162405	9	AC002352	AC002352 Homo sapi	C 779	18	1.5	171867	2	CR628361	CR628361 Danio rer
707	18	1.5	162531	2	AC136231	AC136231 Rattus no	C 780	18	1.5	171887	9	AC024092	AC024092 Homo sapi
708	18	1.5	162712	2	AC015677	AC015677 Homo sapi	C 781	18	1.5	172073	2	AC087307	AC087307 Homo sapi
709	18	1.5	162715	9	AC027569	AC027569 Homo sapi	C 782	18	1.5	172285	2	CR381538	CR381538 Danio rer
710	18	1.5	162715	9	AC108016	AC108016 Homo sapi	C 783	18	1.5	172292	2	AC119269	AC119269 Mus muscu
711	18	1.5	162716	5	BX572632	BX572632 Zebrafish	C 784	18	1.5	172574	2	AC011121	AC011121 Homo sapi
712	18	1.5	162957	2	AC131888	AC131888 Homo sapi	C 785	18	1.5	172585	2	CR847782	CR847782 Danio rer
713	18	1.5	163022	2	AC016851	AC016851 Homo sapi	C 786	18	1.5	172613	9	AC113398	AC113398 Homo sapi
714	18	1.5	163237	9	AC130885	AC130885 Homo sapi	C 787	18	1.5	172965	2	AC023367	AC023367 Homo sapi
715	18	1.5	163283	9	AL355606	AL355606 Human DNA	C 788	18	1.5	173043	9	CNS01DVN	AL135978 Human chr
716	18	1.5	163735	9	AC027277	AC027277 Homo sapi	C 789	18	1.5	173481	9	AC093411	AC093411 Homo sapi
717	18	1.5	163737	9	AC108697	AC108697 Homo sapi	C 790	18	1.5	173580	9	AL445928	AL445928 Human DNA
718	18	1.5	163777	9	AC067932	AC067932 Homo sapi	C 791	18	1.5	174330	2	AC119605	AC119605 Rattus no
719	18	1.5	163932	2	AC146625	AC146625 Papio anu	C 792	18	1.5	174372	2	AC119136	AC119136 Rattus no
720	18	1.5	164000	9	AC121343	AC121343 Homo sapi	C 793	18	1.5	174874	9	AC123595	AC123595 Homo sapi
721	18	1.5	164343	9	AP001099	AP001099 Homo sapi	C 794	18	1.5	175191	9	AC010251	AC010251 Homo sapi
722	18	1.5	164603	2	AC119466	AC119466 Rattus no	C 795	18	1.5	175403	9	AC093268	AC093268 Homo sapi
723	18	1.5	164629	10	AC121094	AC121094 Mus muscu	C 796	18	1.5	175464	2	AP002887	AP002887 Homo sapi
724	18	1.5	164663	4	AC091402	AC091402 Sus scrof	C 797	18	1.5	175583	10	AC125133	AC125133 Mus muscu
725	18	1.5	164744	2	AC018914	AC018914 Homo sapi	C 798	18	1.5	176123	9	AL390718	AL390718 Human DNA
726	18	1.5	164883	9	AC090282	AC090282 Homo sapi	C 799	18	1.5	176258	10	AC147624	AC147624 Mus muscu
727	18	1.5	165097	9	AP003060	AP003060 Homo sapi	C 800	18	1.5	176366	2	AC007941	AC007941 Homo sapi
728	18	1.5	165121	10	AL929312	AL929312 Mouse DNA	C 801	18	1.5	176380	2	AC133461	AC133461 Mus muscu
729	18	1.5	165176	9	AC078941	AC078941 Homo sapi	C 802	18	1.5	176453	2	BX548033	BX548033 Danio rer
730	18	1.5	165774	10	AC102714	AC102714 Mus muscu	C 803	18	1.5	176478	2	AC119237	AC119237 Mus muscu
731	18	1.5	165812	2	AP001141	AP001141 Homo sapi	C 804	18	1.5	176634	2	AC116163	AC116163 Homo sapi
732	18	1.5	165888	2	AC025555	AC025555 Homo sapi	C 805	18	1.5	176940	10	AC132331	AC132331 Mus muscu
733	18	1.5	166063	9	AP002515	AP002515 Homo sapi	C 806	18	1.5	176998	9	AP002846	AP002846 Homo sapi
734	18	1.5	166153	5	BX511140	BX511140 Zebrafish	C 807	18	1.5	177053	2	BX927411	BX927411 Danio rer
735	18	1.5	166229	2	AC022070	AC022070 Homo sapi	C 808	18	1.5	177298	2	AC068751	AC068751 Homo sapi
736	18	1.5	166343	3	CNS01069	AL139785 Human chr	C 809	18	1.5	177787	2	AC060756	AC060756 Homo sapi
737	18	1.5	166352	3	AC010069	AC010069 Drosophil	C 810	18	1.5	177787	9	AC133961	AC133961 Homo sapi
738	18	1.5	166381	2	AC116943	AC116943 Pan trogl	C 811	18	1.5	177894	2	AC117410	AC117410 Homo sapi
739	18	1.5	166464	2	AC148566	AC148566 Canis fam	C 812	18	1.5	177909	2	AC031997	AC031997 Homo sapi
740	18	1.5	166549	2	AC025990	AC025990 Homo sapi	C 813	18	1.5	177929	2	BX942813	BX942813 Danio rer
741	18	1.5	166821	9	AC109582	AC109582 Homo sapi	C 814	18	1.5	178014	9	AC118562	AC118562 Homo sapi
742	18	1.5	166945	2	BX901896	BX901896 Danio rer	C 815	18	1.5	178023	10	AC130718	AC130718 Mus muscu
743	18	1.5	166960	2	AC015958	AC015958 Homo sapi	C 816	18	1.5	178218	2	AC016841	AC016841 Homo sapi
744	18	1.5	167002	2	AC073041	AC073041 Homo sapi	C 817	18	1.5	178242	9	AC008151	AC008151 Homo sapi
745	18	1.5	167102	9	AC094017	AC094017 Homo sapi	C 818	18	1.5	178560	2	AC134877	AC134877 Homo sapi
746	18	1.5	167455	9	AC083281	AC083281 Homo sapi	C 819	18	1.5	178831	5	BX511212	BX511212 Zebrafish
747	18	1.5	167501	2	AC134010	AC134010 Rattus no	C 820	18	1.5	178981	10	AL806511	AL806511 Mouse DNA
748	18	1.5	167540	2	AC016084	AC016084 Homo sapi	C 821	18	1.5	179059	2	AC139303	AC139303 Cercopith
749	18	1.5	167579	9	AC026161	AC026161 Homo sapi	C 822	18	1.5	179357	9	AC007378	AC007378 Homo sapi

823	18	1.5	179391	2	AP001283	AP001283 Homo sapi	c 896	18	1.5	191046	10	AL607146	AL607146 Mouse DNA
824	18	1.5	179532	9	AC107304	AC107304 Homo sapi	c 897	18	1.5	191220	2	AC149647	AC149647 Bos tauru
825	18	1.5	179640	9	AC026462	AC026462 Homo sapi	c 898	18	1.5	191378	2	AC129960	AC129960 Bos tauru
826	18	1.5	179712	3	CNS05TC4	AL355052 Human chr	c 899	18	1.5	191554	2	CR450751	CR450751 Danio rer
827	18	1.5	179869	3	AC104410	AC104410 Drosophi	c 900	18	1.5	191605	2	AC102197	AC102197 Mus muscu
828	18	1.5	180001	9	AC112196	AC112196 Homo sapi	c 901	18	1.5	191765	2	AL392167	AL392167 Human DNA
829	18	1.5	180192	9	AC058810	AC058810 Homo sapi	c 902	18	1.5	191771	2	CR318606	CR318606 Danio rer
830	18	1.5	180194	9	AC012152	AC012152 Homo sapi	c 903	18	1.5	191784	2	AP001774	AP001774 Homo sapi
831	18	1.5	180316	2	AC060786	AC060786 Homo sapi	c 904	18	1.5	192227	10	AC093176	AC093176 Mus muscu
832	18	1.5	180609	9	AC011439	AC011439 Homo sapi	c 905	18	1.5	192297	9	AC040173	AC040173 Homo sapi
833	18	1.5	180759	10	AC129975	AC129975 Mus muscu	c 906	18	1.5	193015	2	AC150074	AC150074 Gallus ga
834	18	1.5	180946	3	AC007827	AC007827 Drosophi	c 907	18	1.5	193188	9	AC092140	AC092140 Homo sapi
835	18	1.5	181058	2	AC068118	AC068118 Homo sapi	c 908	18	1.5	193729	4	AC129099	AC129099 Canis fam
836	18	1.5	181124	9	AC093883	AC093883 Homo sapi	c 909	18	1.5	194063	2	AC118643	AC118643 Mus muscu
837	18	1.5	181336	5	EX088591	EX088591 Zebrafish	c 910	18	1.5	194156	2	AC009851	AC009851 Homo sapi
838	18	1.5	181477	2	AC117799	AC117799 Mus muscu	c 911	18	1.5	194943	9	AC105252	AC105252 Homo sapi
839	18	1.5	181742	10	AC122397	AC122397 Mus muscu	c 912	18	1.5	195375	10	AC119897	AC119897 Mus muscu
840	18	1.5	181936	2	AL360090	AL360090 Homo sapi	c 913	18	1.5	195959	2	AC126692	AC126692 Mus muscu
841	18	1.5	181974	3	AC023724	AC023724 Drosophi	c 914	18	1.5	196115	2	AC148460	AC148460 Xenopus t
842	18	1.5	182032	10	AC122300	AC122300 Mus muscu	c 915	18	1.5	196124	2	AC135684	AC135684 Rattus no
843	18	1.5	182105	3	AC091499	AC091499 Drosophi	c 916	18	1.5	196259	2	AL135915	AL135915 Homo sapi
844	18	1.5	182254	2	AC080139	AC080139 Mus muscu	c 917	18	1.5	196441	2	AC116941	AC116941 Pan trogl
845	18	1.5	182440	9	AC096755	AC096755 Homo sapi	c 918	18	1.5	197757	2	AC074235	AC074235 Homo sapi
846	18	1.5	182628	2	AC015932	AC015932 Mus muscu	c 919	18	1.5	197949	10	AC122204	AC122204 Mus muscu
847	18	1.5	182641	2	AC115885	AC115885 Mus muscu	c 920	18	1.5	197992	9	AC092124	AC092124 Homo sapi
848	18	1.5	182658	2	AC135844	AC135844 Felis cat	c 921	18	1.5	198253	9	AC009090	AC009090 Homo sapi
849	18	1.5	182753	2	AL137253	AL137253 Homo sapi	c 922	18	1.5	198615	10	AC123678	AC123678 Mus muscu
850	18	1.5	182823	9	AC016638	AC016638 Homo sapi	c 923	18	1.5	199080	2	AC122732	AC122732 Mus muscu
851	18	1.5	182935	2	AC142131	AC142131 Rattus no	c 924	18	1.5	199130	10	AL805919	AL805919 Mouse DNA
852	18	1.5	182960	3	AC010017	AC010017 Drosophi	c 925	18	1.5	199130	10	AC099713	AC099713 Mus muscu
853	18	1.5	183062	6	Q0870476	Q0870476 Sequence	c 926	18	1.5	199510	2	AC103122	AC103122 Rattus no
854	18	1.5	183270	2	AC135532	AC135532 Rattus no	c 927	18	1.5	199594	2	AC125284	AC125284 Mus muscu
855	18	1.5	183394	9	AC009102	AC009102 Homo sapi	c 928	18	1.5	199831	5	EX470099	EX470099 Zebrafish
856	18	1.5	183397	10	AL592225	AL592225 Mouse DNA	c 929	18	1.5	200058	2	AC084378	AC084378 Homo sapi
857	18	1.5	183467	2	AC117320	AC117320 Rattus no	c 930	18	1.5	200398	9	AC025566	AC025566 Homo sapi
858	18	1.5	183936	2	AC149620	AC149620 Papio anu	c 931	18	1.5	200509	2	AC006910	AC006910 Caenorhab
859	18	1.5	184065	9	AC023481	AC023481 Homo sapi	c 932	18	1.5	200562	2	AC145859	AC145859 Pan trogl
860	18	1.5	184267	2	AC078796	AC078796 Homo sapi	c 933	18	1.5	200824	2	AC015725	AC015725 Mus muscu
861	18	1.5	184408	8	AP006753	AP006753 Oryza sat	c 934	18	1.5	201493	9	AL691520	AL691520 Human DNA
862	18	1.5	185257	9	AC092704	AC092704 Homo sapi	c 935	18	1.5	201610	5	AL732430	AL732430 Mouse DNA
863	18	1.5	185314	2	AC120244	AC120244 Rattus no	c 936	18	1.5	201659	5	AL953881	AL953881 Zebrafish
864	18	1.5	185399	2	AC026940	AC026940 Homo sapi	c 937	18	1.5	202324	2	AC144712	AC144712 Danio rer
865	18	1.5	185671	5	AC084193	AC084193 Homo sapi	c 938	18	1.5	202377	2	AC144942	AC144942 Mus muscu
866	18	1.5	185760	5	EX296527	EX296527 Zebrafish	c 939	18	1.5	202559	10	AL645849	AL645849 Mouse DNA
867	18	1.5	185726	2	AC151103	AC151103 Bos tauru	c 940	18	1.5	203034	2	AC151572	AC151572 Mus muscu
868	18	1.5	185759	2	AC149612	AC149612 Ictalurus	c 941	18	1.5	203155	2	CR759891	CR759891 Danio rer
869	18	1.5	186127	9	AC104582	AC104582 Homo sapi	c 942	18	1.5	203229	9	AC092718	AC092718 Homo sapi
870	18	1.5	186370	9	AC093671	AC093671 Homo sapi	c 943	18	1.5	203270	9	AC064807	AC064807 Homo sapi
871	18	1.5	186437	2	AC131918	AC131918 Mus muscu	c 944	18	1.5	203450	2	AP002009	AP002009 Homo sapi
872	18	1.5	186635	2	AC051645	AC051645 Homo sapi	c 945	18	1.5	203593	10	AL663053	AL663053 Mouse DNA
873	18	1.5	186908	2	AC023770	AC023770 Homo sapi	c 946	18	1.5	203844	9	AC099667	AC099667 Homo sapi
874	18	1.5	187197	2	AC117766	AC117766 Mus muscu	c 947	18	1.5	204018	2	AC016672	AC016672 Homo sapi
875	18	1.5	187204	2	AC084082	AC084082 Homo sapi	c 948	18	1.5	204508	10	AC098711	AC098711 Mus muscu
876	18	1.5	187266	2	AC073984	AC073984 Homo sapi	c 949	18	1.5	204512	2	AC119029	AC119029 Rattus no
877	18	1.5	187292	2	AC069076	AC069076 Homo sapi	c 950	18	1.5	204600	2	AL683808	AL683808 Mus muscu
878	18	1.5	187393	2	AC084325	AC084325 Mus muscu	c 951	18	1.5	204674	9	AC090164	AC090164 Homo sapi
879	18	1.5	187456	2	AC036130	AC036130 Homo sapi	c 952	18	1.5	204777	9	AP001011	AP001011 Homo sapi
880	18	1.5	187540	2	AC073970	AC073970 Homo sapi	c 953	18	1.5	204859	2	EX901944	EX901944 Danio rer
881	18	1.5	187634	10	AC125110	AC125110 Mus muscu	c 954	18	1.5	205056	2	AC132734	AC132734 Rattus no
882	18	1.5	187962	2	CR759825	CR759825 Danio rer	c 955	18	1.5	205588	9	AC068400	AC068400 Homo sapi
883	18	1.5	188337	2	CEY1787	Z92850 Caenorhabdi	c 956	18	1.5	206472	2	AC126091	AC126091 Rattus no
884	18	1.5	188414	2	CR388091	CR388091 Danio rer	c 957	18	1.5	207475	10	AC126032	AC126032 Mus muscu
885	18	1.5	188557	2	AC106737	AC106737 Homo sapi	c 958	18	1.5	207835	9	AC019188	AC019188 Homo sapi
886	18	1.5	188757	2	AC007903	AC007903 Homo sapi	c 959	18	1.5	208035	2	AC134450	AC134450 Mus muscu
887	18	1.5	188765	2	AC145051	AC145051 Canis fam	c 960	18	1.5	208198	10	AC107664	AC107664 Mus muscu
888	18	1.5	189125	2	CR392339	CR392339 Danio rer	c 961	18	1.5	208670	9	AC103588	AC103588 Homo sapi
889	18	1.5	189181	2	AP001333	AP001333 Homo sapi	c 962	18	1.5	208869	2	AC097990	AC097990 Rattus no
890	18	1.5	189189	2	AC128605	AC128605 Rattus no	c 963	18	1.5	208921	2	AP001911	AP001911 Homo sapi
891	18	1.5	189768	2	CR753421	CR753421 Homo sapi	c 964	18	1.5	208944	10	AC131065	AC131065 Mus muscu
892	18	1.5	189768	2	AC044889	AC044889 Homo sapi	c 965	18	1.5	209123	9	AC100812	AC100812 Homo sapi
893	18	1.5	189771	2	AC112921	AC112921 Homo sapi	c 966	18	1.5	209586	2	AC148497	AC148497 Otolenur
894	18	1.5	190388	2	AC108439	AC108439 Mus muscu	c 967	18	1.5	209714	10	AL806532	AL806532 Mouse DNA
895	18	1.5	190998	5	EX649384	EX649384 Zebrafish	c 968	18	1.5	209764	10	AC133646	AC133646 Mus muscu

c 969	18	1.5	209810	2	AC106935	AC106935 Rattus no
c 970	18	1.5	209888	2	AC146462	AC146462 Saimiri s
c 971	18	1.5	210220	2	AC127868	AC127868 Rattus no
c 972	18	1.5	210359	9	HSR430K20	AL121898 Human DNA
c 973	18	1.5	210376	2	AL136118	AL136118 Homo sapi
c 974	18	1.5	210976	2	AC142430	AC142430 Rattus no
c 975	18	1.5	211515	9	AC144479	AC144479 Pan trogl
c 976	18	1.5	211654	2	AL591103	AL591103 Homo sapi
c 977	18	1.5	212421	2	AC109591	AC109591 Homo sapi
c 978	18	1.5	212611	2	AC115143	AC115143 Rattus no
c 979	18	1.5	213131	5	EX663611	EX663611 Zebrafish
c 980	18	1.5	213393	2	AL645475	AL645475 Homo sapi
c 981	18	1.5	213409	10	AC113318	AC113318 Mus muscu
c 982	18	1.5	213746	2	CR847568	CR847568 Danio rer
c 983	18	1.5	214014	2	AC109845	AC109845 Rattus no
c 984	18	1.5	214140	2	EX663604	EX663604 Danio rer
c 985	18	1.5	214181	2	AC137479	AC137479 Rattus no
c 986	18	1.5	214540	2	AC111834	AC111834 Rattus no
c 987	18	1.5	214836	2	AC093869	AC093869 Homo sapi
c 988	18	1.5	214945	9	AC098588	AC098588 Homo sapi
c 989	18	1.5	215087	2	AC113594	AC113594 Mus muscu
c 990	18	1.5	215147	10	AC119806	AC119806 Mus muscu
c 991	18	1.5	215339	2	AC115427	AC115427 Rattus no
c 992	18	1.5	215480	2	AC135781	AC135781 Homo sapi
c 993	18	1.5	215724	2	AC099902	AC099902 Mus muscu
c 994	18	1.5	215861	2	AC102408	AC102408 Mus muscu
c 995	18	1.5	215910	2	EX928752	EX928752 Danio rer
c 996	18	1.5	216082	10	AC107369	AC107369 Mus muscu
c 997	18	1.5	216362	10	AC099692	AC099692 Mus muscu
c 998	18	1.5	216968	2	EX649391	EX649391 Danio rer
c 999	18	1.5	216969	2	AC132185	AC132185 Rattus no
1000	18	1.5	217746	9	AC015909	AC015909 Homo sapi

ALIGNMENTS

RESULT 1	BD205212	1181 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).				
DEFINITION	BD205212				
ACCESSION	BD205212.1 GI:33014982				
VERSION	JP 2002512813-A/2.				
KEYWORDS	unidentified				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1181)				
AUTHORS	Frechon,D.T.M., Laure,F.C. and Thierry,D.				
TITLE	Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)				
JOURNAL	Patent: JP 2002512813-A 2 08-MAY-2002;				
COMMENT	BIORAD PASTEUR				
OS	Unidentified				
	PN JP 2002512813-A/2				
	PD 08-MAY-2002				
	PF 27-APR-1999 JP 2000546051				
	PR 28-APR-1998 FR 98/05329				
FI	DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDE LAURE, PI				
	DOMINIQUE THIERRY				
	PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC				
	Strandedness: Double;				
	CC Topology: Linear;				
CC	Nucleotide sequence for detecting enterohemorrhagic CC				
	(EHEC).				
	Escherichia coli				
	Key				
	Location/Qualifiers				
FT	1..1181				
	source				
	Location/Qualifiers				
	/organism='Unidentified'.				
	1..1181				
FEATURES	/organism="unidentified"				
	source				

ORIGIN		/mol_type="genomic DNA"		/db_xref="taxon:32644"	
Query Match		100.0%; Score 1181; DB 6; Length 1181;			
Best Local Similarity		100.0%; Pred. No. 0;			
Matches 1181; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CTGCAGGAGATGAAAAAGCCAAATAAAAAATTTGCCATCCACGCGCGCTCCAGCTG	60		
Db	1	CTGCAGGAGATGAAAAAGCCAAATAAAAAATTTGCCATCCACGCGCGCTCCAGCTG	60		
Qy	61	AAAGTAGGCCCTGTTCTGTCGGGTATTTAAATGATTTGAACCGTCCCGTATTTAAACAATG	120		
Db	61	AAAGTAGGCCCTGTTCTGTCGGGTATTTAAATGATTTGAACCGTCCCGTATTTAAACAATG	120		
Qy	121	TGATAAATTTACTCCGTTTACCGGAAAACCGCTGAACAAAATTTCCGGCTGAAAAGAGATCC	180		
Db	121	TGATAAATTTACTCCGTTTACCGGAAAACCGCTGAACAAAATTTCCGGCTGAAAAGAGATCC	180		
Qy	181	GCGGTTATCTGTTGTCATTTCCCTTAGACCTGACTAGCCAGAGACACAATGATCTGTGCCG	240		
Db	181	GCGGTTATCTGTTGTCATTTCCCTTAGACCTGACTAGCCAGAGACACAATGATCTGTGCCG	240		
Qy	241	TTCTGTTAATATCAAAACCGGTACTCAATATCTTCTTGGCGCTGGCTGCATCATCCGGA	300		
Db	241	TTCTGTTAATATCAAAACCGGTACTCAATATCTTCTTGGCGCTGGCTGCATCATCCGGA	300		
Qy	301	AGCGTTCCGCTCCGGATAAAAAATTCGGCAGTCGCGCGCTCATGACAGACACATCCCCCA	360		
Db	301	AGCGTTCCGCTCCGGATAAAAAATTCGGCAGTCGCGCGCTCATGACAGACACATCCCCCA	360		
Qy	361	CGGGTAAACAGCGTCCCTGTACATCTTCTTGAATGACATCAGGGATCCCGCCCGTCTCAC	420		
Db	361	CGGGTAAACAGCGTCCCTGTACATCTTCTTGAATGACATCAGGGATCCCGCCCGTCTCAC	420		
Qy	421	TGGCGATAACCGGCAACCGCGAGACTGACGCTTTCAGCCAGTACCATACCAACGCTTCAT	480		
Db	421	TGGCGATAACCGGCAACCGCGAGACTGACGCTTTCAGCCAGTACCATACCAACGCTTCAT	480		
Qy	481	TTTTCGGAAGGCATGACCAACACACTGGGCAATCCGGTAGACCGGTAAACGCTGGGAAAAGGG	540		
Db	481	TTTTCGGAAGGCATGACCAACACACTGGGCAATCCGGTAGACCGGTAAACGCTGGGAAAAGGG	540		
Qy	541	CACCTGCCATTAAACACATCTCGCTCATTTCCAGGTGTTCTGTCTGCTACGCGACGACGTG	600		
Db	541	CACCTGCCATTAAACACATCTCGCTCATTTCCAGGTGTTCTGTCTGCTACGCGACGACGTG	600		
Qy	601	CTTTCGTATTTCTTCAGCCCGCGGCCACCAACGAGCCAGCGAAATGATTTCCCTTCCATCT	660		
Db	601	CTTTCGTATTTCTTCAGCCCGCGGCCACCAACGAGCCAGCGAAATGATTTCCCTTCCATCT	660		
Qy	661	TCAGCTGATACAAATACACGACGATAAATTCATGCTCTTTTCGGGACGTAGCATCCCCA	720		
Db	661	TCAGCTGATACAAATACACGACGATAAATTCATGCTCTTTTCGGGACGTAGCATCCCCA	720		
Qy	721	CCGTGAACGATAACGGGAAACATTGCTGTGTGATGTCAGCCCGGCTGGATATGACAGGGGTA	780		
Db	721	CCGTGAACGATAACGGGAAACATTGCTGTGTGATGTCAGCCCGGCTGGATATGACAGGGGTA	780		
Qy	781	ACGGTCCGATGGCTTCATTATGCAATTCGCGGCCAGTCGAAACCCCGGTGGAATAACCGTTA	840		
Db	781	ACGGTCCGATGGCTTCATTATGCAATTCGCGGCCAGTCGAAACCCCGGTGGAATAACCGTTA	840		
Qy	841	CCGGTGTCTGACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGGCACAA	900		
Db	841	CCGGTGTCTGACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGGCACAA	900		
Qy	901	TGAAATTCACACAGATAAATTCAGGAAAACGTTCTGTGCTTACGGGTGATGATGTTTTTT	960		
Db	901	TGAAATTCACACAGATAAATTCAGGAAAACGTTCTGTGCTTACGGGTGATGATGTTTTTT	960		
Qy	961	GTCTGACAAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCTGCTATATTACTGT	1020		

Db	961	GTCTGACAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCCTGCCTATATTACTGT	1020
Qy	1021	CATGGCCACTATGCGCAGATGACCAAGATCAGGTTTAAATTCGCCGATTAATCCGTCGAAGTC	1080
Db	1021	CATGGCCACTATGCGCAGATGACCAAGATCAGGTTTAAATTCGCCGATTAATCCGTCGAAGTC	1080
Qy	1081	TGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGCAATAAAGTCACATCATGCCCTC	1140
Db	1081	TGAGGATGGAAGGAAGGTGAAGGCTGTTCTCTGAAAGCAATAAAGTCACATCATGCCCTC	1140
Qy	1141	TTTTTCTGGCTTCGGAGCAATTTATCTTTTCTCTGCAG	1181
Db	1141	TTTTTCTGGCTTCGGAGCAATTTATCTTTTCTCTGCAG	1181

RESULT 2					
AX011298					
LOCUS	AX011298	1181 bp	DNA	linear	PAT 06-SEP-2000
DEFINITION	Sequence 2 from Patent WO9955908.				
ACCESSION	AX011298				
VERSION	AX011298.1	GI:9997848			

SOURCE	ORGANISM
Escherichia coli	Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia.	Escherichia coli

REFERENCE	1
AUTHORS	Thierry, D., Frechon, D.T. and Laure, F.C.
TITLE	Nucleotide sequences for detecting enterohemorrhagic <i>escherichia coli</i> (ehc)
JOURNAL	Patent: WO 955908-A 2 04-NOV-1999; THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE FRANÇOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)

```
FEATURES
    source
    1. .1181
    Location/Qualifiers
        /organism="Escherichia coli"
        /mol_type="unassigned DNA"
        /db_xref="taxon:562"
```

Query Match	100.0%	Score 1181;	DB 6;	Length 1181;
Best Local Similarity	100.0%;	Fred. No. 0;		
Matches 1181;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy	1	CTG	CAG	GAT	TG	AAAA	AAAA	AGC	CAAA	ATAA	AAAA	TTG	CCCA	TCCC	CAG	CGCG	TCC	CAG	TG	60
Db	1	CTG	CAG	GAT	TG	AAAA <td>AAAA</td> <td>AGC</td> <td>CAAA</td> <td>ATAA</td> <td>AAAA</td> <td>TTG</td> <td>CCCA</td> <td>TCCC</td> <td>CAG</td> <td>CGCG <td>TCC</td> <td>CAG</td> <td>TG</td> <td>60</td> </td>	AAAA	AGC	CAAA	ATAA	AAAA	TTG	CCCA	TCCC	CAG	CGCG <td>TCC</td> <td>CAG</td> <td>TG</td> <td>60</td>	TCC	CAG	TG	60

Qy	61	AAAGTAGCCCTGTTCTGTGCCGGTATTTAAATGCATTCGATCCGTCCTCCCGTATTAAACAATG	120
Db	61 <td>AAAGTAGCCCTGTTCTGTGCCGGTATTTAAATGCATTCGATCCGTCCTCCCGTATTAAACAATG <td>120</td> </td>	AAAGTAGCCCTGTTCTGTGCCGGTATTTAAATGCATTCGATCCGTCCTCCCGTATTAAACAATG <td>120</td>	120

Qy	121	TGATAAATTACTCCGTTACCGGAAACCGCTGAACAAATTCGGGCTGAAAGAGGATCC	180
Db	121 <th>TGATAAATTACTCCGTTACCGGAAACCGCTGAACAAATTCGGGCTGAAAGAGGATCC</th> <th>180</th>	TGATAAATTACTCCGTTACCGGAAACCGCTGAACAAATTCGGGCTGAAAGAGGATCC	180

Qy	181	GC	GT	TAT	CT	GT	GC	ATT	CC	CC	TAG	CCT	GACT	AG	CC	AG	AG	CA	CA	AT	TG	AT	CT	GT	GC	CG	240
Dy <td>181 <td>GC <td>GT <td>TAT <td>CT <td>GT <td>GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	181 <td>GC <td>GT <td>TAT <td>CT <td>GT <td>GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GC <td>GT <td>TAT <td>CT <td>GT <td>GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GT <td>TAT <td>CT <td>GT <td>GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TAT <td>CT <td>GT <td>GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CT <td>GT <td>GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GT <td>GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GC <td>ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	ATT <td>CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	TAG <td>CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CCT <td>GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GACT <td>AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	AG <td>CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td></td>	CC <td>AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td></td>	AG <td>AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td></td>	AG <td>CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td></td>	CA <td>CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td></td>	CA <td>AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td></td>	AT <td>TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td></td>	TG <td>AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td></td>	AT <td>CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td></td>	CT <td>GT <td>GC <td>CG <td>240</td> </td></td></td>	GT <td>GC <td>CG <td>240</td> </td></td>	GC <td>CG <td>240</td> </td>	CG <td>240</td>	240

[illegible]

Qy	301	AGCGTTCCGGTGGGATAAAAAATCGCGCAGTGC	CGCGGTCCATGCAGACACATCCGCCA	360
Dh	301	AGCGTTCCGGTGGGATAAAAAATCGCGCAGTGC	CGCGGTCCATGCAGACACATCCGCCA	360

QY	361	CGGGTAACAGCGTCCCTGTACACATTTCTCTGAATGACATCAGGGATCCGCGCGGTCTCAC	420
DB	361	CGGGTATTCAGCGTCCCTGTACACATTTCTCTGAATGACATCAGGGATCCGCGCGGTCTCAC	420

Qy	421	TGGCGATAACGGGCA	CGCCGGAGCTGA	CGCTTCAGCAGTACCAT	CCAAACGCTTCAT	480
Db	421	TGGCGATAACGGGCA	CGCCGGAGCTGA	CGCTTCAGCAGTACCAT	CCAAACGCTTCAT	480
Qy	481	TTTCCGAAGGCATG	ACCAACACACTG	CGAATCCGGTAGAC	CCGGTAAACGCTGGGAAAGGG	540
Db	481	TTTCCGAAGGCATG	ACCAACACACTG	CGAATCCGGTAGAC	CCGGTAAACGCTGGGAAAGGG	540
Qy	541	CACCTGCCATTAA	CACATCTCCGCTCAT	TCTCCAGGTTCTGTCTG	CTGTGACGACGACGTG	600
Db	541	CACCTGCCATTAA	CACATCTCCGCTCAT	TCTCCAGGTTCTGTCTG	CTGTGACGACGACGTG	600
Qy	601	CTTTCGTATTC	TTTCAAGCCGGGCGCC	CACACGACGACCGA	ATGATTTCCCTTCATCT	660
Db	601	CTTTCGTATTC	TTTCAAGCCGGGCGCC	CACACGACGACCGA	ATGATTTCCCTTCATCT	660
Qy	661	TCAGCTGATACA	ATACACGACGATAA	ATTCATGTCTCTTTTC	CGGACGCTAGCATCCCA	720
Db	661	TCAGCTGATACA	ATACACGACGATAA	ATTCATGTCTCTTTTC	CGGACGCTAGCATCCCA	720
Qy	721	CCTGAAACGATP	AACGGGAACATTTCT	CTGTAACGACCCACGCGT	GGATATGACGGGTA	780
Db	721	CCTGAAACGATP	AACGGGAACATTTCT	CTGTAACGACCCACGCGT	GGATATGACGGGTA	780
Qy	781	ACGGTGCATGGCTT	CATTATGCAATCGGGCCAGT	CGAAACCCGGTGGNAAT	AACCGTTA	840
Db	781	ACGGTGCATGGCTT	CATTATGCAATCGGGCCAGT	CGAAACCCGGTGGNAAT	AACCGTTA	840
Qy	841	CCGGTGTCTGAC	ACCTTCGCCATCAGAT	TGCGCCATCATGGGTG	AGATAGGCACAA	900
Db	841	CCGGTGTCTGAC	ACCTTCGCCATCAGAT	TGCGCCATCATGGGTG	AGATAGGCACAA	900
Qy	901	TGAATACACAGAT	AAATTCAGGGAACCGT	TCTTGCTTTACGGGTG	ATGTAGGTTTTT	960
Db	901	TGAATACACAGAT	AAATTCAGGGAACCGT	TCTTGCTTTACGGGTG	ATGTAGGTTTTT	960
Qy	961	GTCTGACATATG	TGAAGCGGTGACAGCAT	ATCAGACGGCTCAGT	CTCTGCTATATTACTGT	1020
Db	961	GTCTGACATATG	TGAAGCGGTGACAGCAT	ATCAGACGGCTCAGT	CTCTGCTATATTACTGT	1020
Qy	1021	CATGGCCATCTAT	GGCAGATGACAGAT	CAGGTTTAAATTC	CCCGATAATCCGTCGAGTC	1080
Db	1021	CATGGCCATCTAT	GGCAGATGACAGAT	CAGGTTTAAATTC	CCCGATAATCCGTCGAGTC	1080
Qy	1081	TGAGGATGGAAG	GGAAGGTGAAGGCTGT	TCTGAAAGGAAT	ATAAAGTGACATCATG	1140
Db	1081	TGAGGATGGAAG	GGAAGGTGAAGGCTGT	TCTGAAAGGAAT	ATAAAGTGACATCATG	1140
Qy	1141	TTTTTCTGGCTT	CCGAGCAATTTTACT	TTTTTCTCTGAG	1181	
Db	1141	TTTTTCTGGCTT	CCGAGCAATTTTACT	TTTTTCTCTGAG	1181	

[illegible]

COMMENT FEATURES source	Related Location/Qualifiers 1..7395 /organism="Escherichia coli" /mol_type="genomic DNA" /strain="EDL933" /serotype="O157:H7" /db_xref="taxon:562" /clone="pSm7" /plasmid="pO157" 774..1595 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAA72138.1" /db_xref="GI:4127813" /db_xref="GOA:O85100" /db_xref="UniProt/TREMBL:O85100" /translations="MLNARHLPVLMVHHVSRCPGLVLTSPETFEQMKWLADNNWRTV TSABLYFGCGTLPKRSVMLTFDDGVLNWLRAWPLVKEYNLHAHIFLITGLIGKN VRSRQEVSHRDCERLIADNRSDVLMRWSEVREMDSLGFVHFVHTSHKSRWRLS VSRQCRMLKEDILVGKQCLTEKLGFCSSHLCPGEGYNRDIYNLAGLHGFSLYIT ERRNCPENSGLRIGRISTKERESHGWLKRRLFYITPLFSSVLALHKGRLPDN" 1595..2701 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAA72139.1" /db_xref="GI:4127814" /db_xref="GOA:O82935" /db_xref="UniProt/TREMBL:O82935" /translations="MKKILPTSSDIGOGLQALQAMTALQKHSHVLIACREKSKI APARKGHQVTFIPPRNSLHPLSLRLRIIGEFKPDLVICHSGHDSNIAGLSRLIC CHFSIVRQKTYITRKTFTSLANVCDPIVPSANMAHLMASGVRTPVTVIIPGDM PALHNEARPLPHIHAWAADNPLVIVQGMREKGFHFMRLVLIQLKMEGKSPR WLVVAGREYERARLQQTQHEHGMGSDVLMAGALFPALPVYIASVVMPSSEAFGM VLAEVSGVPEYVASETGGTIPVIOKNVTGTLPLVGDVSAMTGAIRDFLSRPRFRM AASAREDIERYFDINRTAQIIVSLASQAKGKCNR" 2791..4512 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAA72140.1" /db_xref="GI:4127815" /db_xref="GOA:O9R768" /db_xref="UniProt/TREMBL:O9R768" /translations="MLNLTQGNRPTFSWALGWAIFVFGFFSTLLQVIFPSSGYSGTN GIRDSLFLSCLVLPVLYPDRIKIIAAVVGFIWLGTSLAALCYFLYGHFSSQVLF VMPETNAREAGEYFSQVSLKLLISLVYTAVSFVLTWRLRPVYIPLPWRRIIVSPLL YALLHPVLKSLIROBPLNDTLGKLASMEPAAPWOFVSSYYQYHQNALNLTFLNE NSALPLGNLRDSSGERPRTLVLVIGESTORERMSLYGYLRETTPELDALARKTDGILT VFNNVASRPYTIIEALQOALTANENKPDILYLTQPSLMMNMKQAGYKTFWITNQOIT ARNMLTVRQRDRQTYMNRQRTQSAREYDYNVLPFFREVLNDPAPKLLIIVHLLGT HIKRYKYPGQRFQDGTIGHIPTGLNAKEYNDYNDNANFVNDVAVSLIKDFRAT APDGFLLYDHGEEVYDTPPYKTOGRNEDNTPRMVYVPLFWLTSEKMHAAHPDFS QYVDRKYSLSAEHLHTWSDLAGLVYDGVDPTRSLVNPQFRETTRWIGNPYKKNGLTDFD TLPVGEF" 4586..5584 /gene="mebB" 4586..5584 /gene="mebB" /codon_start=1 /transl_table=11 /product="MebB protein" /protein_id="CAA72141.1" /db_xref="GI:4127816" /db_xref="GOA:O82937" /db_xref="UniProt/TREMBL:O82937" /translations="MVPSPAVLCVHNEISRQIPVNMKNIRTEFIPRNLTLCPRYWM TWNGIGICVFAMVPALRDLPLGLKSGARQALINLSICPEYSDKEKE NIYDAMPATASMAVLMALSGPDKISHRIRWNGLEIVKMAQONNEKVIPLVPHAW GVDPAPMLMAASARMAAFHNPVVDIYWNVSRFRFGKLUHARDNGIASVRSVR QGYWGYLPPQDHGPFSEFADFATYKATLPVIGRLSRISGARIIPLFPVYDGRKTHH
CDS	misc_feature 5645..5820 /note="IS3-like sequence" 6465..>7395 /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="CAA72142.1" /db_xref="GI:4127817" /db_xref="GOA:O9ZAL1" /db_xref="UniProt/TREMBL:O9ZAL1" /translations="MNTQNERWRTPMKLYLSCTTILAPLAIGVFSATAADNNSAIYF NTSQPINDLQGLSAAEVKFAQSQILPAHFKEGDSQPHLSLRKSLLLVRPVKADKTP VLSAARDNNKILGTLTLYPPSLPDTIIVHLDGVPBGIGDFTPHNGTKKIINTVAEVN KLSDSGSSISHLHTNNALVEIHTANGRWVDIYLPQGDLEGKMFVFTSSAGVSTV FYGDRKVTLSVGNLTLLPKYVNGQWFRSGELENNRITYAQHINSAEI PAHWIVPGLNLV IKQGNLSGRINDIKIGAPGELLHTIDIGMLTTP"
CDS	ORIGIN Query Match 86.2%; Score 1018; DB 1; Length 7395; Best Local Similarity 99.7%; Pred. No. 0; Matches 1188; Conservative 0; Mismatches 3; Indels 0; Gaps 0; QY 11 TGGAAAAAGGCCAAATAATAAAATTTGCCCATCCACGCGCTCCAGCTGAAAGTAGGCC 70 DB 2875 TGGAAAAAGGCCAAATAATAAAATTTGCCCATCCACGCGCTCCAGCTGAAAGTAGGCC 2816 QY 71 TGTTCCTGTCGGTATTAAATGCAATTGACCGTCCCGTATTAAACAATGTGATAATTA 130 DB 2815 TGTTCCTGTCGGTATTAAATGCAATTGACCGTCCCGTATTAAACAATGTGATAATTA 2756 QY 131 CTCGGTTACCGAAAAACCGCTGAACAAAAATTCGGGCTGAAAAGAGATCCGCCGTTATCT 190 DB 2755 CTCGGTTACCGAAAAACCGCTGAACAAAAATTCGGGCTGAAAAGAGATCCGCCGTTATCT 2696 QY 191 GTTGCAATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTGTAAT 250 DB 2695 GTTGCAATTTCCCTTAGCCTGACTAGCCAGAGACACAATGATCTGTGCCGTTCTGTGTAAT 2636 QY 251 ATCAAAACCGGTACTCAATATCTTCTGCGGCTGGCTGCCATCATCCGGAAGCGTTCCGG 310 DB 2635 ATCAAAACCGGTACTCAATATCTTCTGCGGCTGGCTGCCATCATCCGGAAGCGTTCCGG 2576 QY 311 TCGGGATAAAAAATCCGCGAGTGCCTGATGACAGACACATCCCCACGCGTAAACAG 370 DB 2575 TCGGGATAAAAAATCCGCGAGTGCCTGATGACAGACACATCCCCACGCGTAAACAG 2516 QY 371 CGTCCCTGTACATTTCTTGAATGACATCAGGGATCCCGCCCGCTCTCACTGGCGATAAC 430 DB 2515 CGTCCCTGTACATTTCTTGAATGACATCAGGGATCCCGCCCGCTCTCACTGGCGATAAC 2456 QY 431 GGGCAGCCGGGAGCTGACGCTTACGCCAGTACCATACCAACGCTTCATTTTCCGAAGG 490 DB 2455 GGGCAGCCGGGAGCTGACGCTTACGCCAGTACCATACCAACGCTTCATTTTCCGAAGG 2396 QY 491 CATGACACACACTCGGCAATCCGCTAGACCGGTAAACGCTGGGAAAAGGCACCTGCCAT 550 DB 2395 CATGACACACACTCGGCAATCCGCTAGACCGGTAAACGCTGGGAAAAGGCACCTGCCAT 2336 QY 551 TAAACATCTCCGCTCATTTCCAGGTGTTCTGTCTGTGACGAGCGTGTCTTCCATCTTCAGCTGATA 610 DB 2335 TAAACATCTCCGCTCATTTCCAGGTGTTCTGTCTGTGACGAGCGTGTCTTCCATCTTCAGCTGATA 2276 QY 611 TTACCCCGCGGCCCCACACGAGCCAGCGAATGATTTCCCTTCATCTTCAGCTGATA 670 DB 2275 TTACCCCGCGGCCCCACACGAGCCAGCGAATGATTTCCCTTCATCTTCAGCTGATA 2216 QY 671 CAATACACGAGCATATAATTCATGTCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 730 DB 2215 CAATACACGAGCATATAATTCATGTCTTTTTCGGGACGTAGCATCCCCACCTGAACGAT 2156 QY 731 AAGCGGAACATTGCTGCTGATGACGCCACCGCTGATATGACGGGGTAACCGGTGCGCAT 790


```
Db 2155 AAGCGGAACATTGTCTCTGATGACGCCAGCGGTGATATGACGGGTAAACGGTCCCAT 2096
Qy 791 GGCTTCATTATGCAATCGCGCCAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGTCTCT 850
Db 2095 GGCTTCATTATGCAATCGCGCCAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGTCTCT 2036
Qy 851 GACACCTTCGGCCATCAGATCGGCCATCATCGGTGAGATGAGGACAAACATGAATCACA 910
Db 2035 GACACCTTCGGCCATCAGATCGGCCATCATCGGTGAGATGAGGACAAACATGAATCACA 1976
Qy 911 CAGATAATTTCAGGGAACCGTTCTGGTCTTACGGGTGATGTAGGTTTTTTTGTCTGACAAT 970
Db 1975 CAGATAATTTCAGGGAACCGTTCTGGTCTTACGGGTGATGTAGGTTTTTTTGTCTGACAAT 1916
Qy 971 AGTGAAGCGGTGACAGATATCAGACGGCTCAGTCTCTGCTATATTAATCTGTGCGCACT 1030
Db 1915 ACTGAAGCGGTGACAGATATCAGACGGCTCAGTCTCTGCTATATTAATCTGTGCGCACT 1856
Qy 1031 ATGCGAGATCAGCAGATCAGGTTTAAATTCGCCGTAATCCGTCGAAGTCTGAGGATGA 1090
Db 1855 ATGCGAGATCAGCAGATCAGGTTTAAATTCGCCGTAATCCGTCGAAGTCTGAGGATGA 1796
Qy 1091 AGGAAGGTGAAGGCTGTCTCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTCTGGC 1150
Db 1795 AGGAAGGTGAAGGCTGTCTCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTCTGGC 1736
Qy 1151 TTCGGAGCAATTTTACTTTTTTCTCTGCGAG 1181
Db 1735 TTCGGAGCAATTTTACTTTTTTCTCTGCGAG 1705

RESULT 4
AX191728/c
LOCUS
DEFINITION
Sequence 10 from Patent WO0149775.
ACCESSION
AX191728
VERSION
AX191728.1 GI:15209897
KEYWORDS
Escherichia coli
SOURCE
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
ORGANISM
Enterobacteriaceae; Escherichia.
REFERENCE
1
AUTHORS
Iversen, P.L.
TITLE
Antisense antibacterial cell division composition and method
JOURNAL
Patent: WO 0149775-A 10 12-JUL-2001;
Avi Biopharma, Inc. (US)
FEATURES
Location/Qualifiers
source
1..7395
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 86.2%; Score 1018; DB 6; Length 7395;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 TGGAAAAAGCCAAATAAAAAATTGCCCATCCAGCGGCTCCAGCTGAAAGTAGGCC 70
Db 2875 TGGAAAAAGCCAAATAAAAAATTGCCCATCCAGCGGCTCCAGCTGAAAGTAGGCC 2816
Qy 71 TGTCTGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
Db 2815 TGTCTGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 2756
Qy 131 CTCGGTTACCGGAACCGCTGACAAATAATTCGGCTGAAAGAGATCCGCGTTATCT 190
Db 2755 CTCGGTTACCGGAACCGCTGACAAATAATTCGGCTGAAAGAGATCCGCGTTATCT 2696
Qy 191 GTTGCAATTTCCCTTACGCTGACTAGCCAGAGACACAATGATCTGTGCGGCTTCTGTTAAT 250
Db 2695 GTTGCAATTTCCCTTACGCTGACTAGCCAGAGACACAATGATCTGTGCGGCTTCTGTTAAT 2636
```

```
Qy 251 ATCAAAACCGGTACTCAATATCTTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCCG 310
Db 2635 ATCAAAACCGGTACTCAATATCTTTCTCTGGCGCTGGCTGCCATCATCCGGAAGCGTTCCCG 2576
Qy 311 TCGGGTAATAAATCGCGCAGTGGCGGCTCCATGCGAGACACATCCCCACCGGTACACAG 370
Db 2575 TCGGGTAATAAATCGCGCAGTGGCGGCTCCATGCGAGACACATCCCCACCGGTACACAG 2516
Qy 371 CGTCCCTGTCAATCTTCTGTAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 430
Db 2515 CGTCCCTGTCAATCTTCTGTAATGACATCAGGGATCCCGCCGCTCTCACTGGCGATAAC 2456
Qy 431 GGGCAGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAAAACGCTTCATTTTCCGAAGG 490
Db 2455 GGGCAGCGCGGAGACTGACGCTTCAGCCAGTACCATACCAAAACGCTTCATTTTCCGAAGG 2396
Qy 491 CATGACCCACACACTGGCAATCCGGTAGACCGGTAAAGTGGGAAAGGGCACTGCCAT 550
Db 2395 CATGACCCACACACTGGCAATCCGGTAGACCGGTAAAGTGGGAAAGGGCACTGCCAT 2336
Qy 551 TAAACATCTCCGCTCATTTCCAGAGTGTCTGTCTGCTGACGAGAGCTGCTTCGTATTC 610
Db 2335 TAAACATCTCCGCTCATTTCCAGAGTGTCTGTCTGCTGACGAGAGCTGCTTCGTATTC 2276
Qy 611 TTACGCCCGCGGCCACCAACGAGCCAGGAAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 2275 TTACGCCCGCGGCCACCAACGAGCCAGGAAATGATTTCCCTTCCATCTTCAGCTGATA 2216
Qy 671 CAATACACCGCAGCATAAATTCATGCTCTTTTCGGGACGTAGCATCCCCACCTGAAAGAT 730
Db 2215 CAATACACCGCAGCATAAATTCATGCTCTTTTCGGGACGTAGCATCCCCACCTGAAAGAT 2156
Qy 731 AAGCGGAACATTGTCTGTCTGATGACGCCAGCGGTGGATATGACGGGTAAACGGTCCCAT 790
Db 2155 AAGCGGAACATTGTCTGTCTGATGACGCCAGCGGTGGATATGACGGGTAAACGGTCCCAT 2096
Qy 791 GGCTTCATTATGCAATCGCGCCAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGTCTCT 850
Db 2095 GGCTTCATTATGCAATCGCGCCAGTCGAAACCCGGTGGAAATAACCGTTACCGGTGTCTCT 2036
Qy 851 GACACCTTCGGCCATCAGATCGGCCATCATCGGTGAGATGAGGACAAACATGAATCACA 910
Db 2035 GACACCTTCGGCCATCAGATCGGCCATCATCGGTGAGATGAGGACAAACATGAATCACA 1976
Qy 911 CAGATAATTTCAGGGAACCGTTCTGGTCTTACGGGTGATGTAGGTTTTTTTGTCTGACAAT 970
Db 1975 CAGATAATTTCAGGGAACCGTTCTGGTCTTACGGGTGATGTAGGTTTTTTTGTCTGACAAT 1916
Qy 971 AGTGAAGCGGTGACAGATATCAGACGGCTCAGTCTCTGCTATATTAATCTGTGCGCACT 1030
Db 1915 ACTGAAGCGGTGACAGATATCAGACGGCTCAGTCTCTGCTATATTAATCTGTGCGCACT 1856
Qy 1031 ATGCGAGATCAGCAGATCAGGTTTAAATTCGCCGTAATCCGTCGAAGTCTGAGGATGA 1090
Db 1855 ATGCGAGATCAGCAGATCAGGTTTAAATTCGCCGTAATCCGTCGAAGTCTGAGGATGA 1796
Qy 1091 AGGAAGGTGAAGGCTGTCTCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTCTGGC 1150
Db 1795 AGGAAGGTGAAGGCTGTCTCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTCTGGC 1736
Qy 1151 TTCGGAGCAATTTTACTTTTTTCTCTGCGAG 1181
Db 1735 TTCGGAGCAATTTTACTTTTTTCTCTGCGAG 1705
```

RESULT 5

AF074613/c

LOCUS

DEFINITION Escherichia coli O157:H7 plasmid pO157, complete sequence.

ACCESSION AF074613

VERSION AF074613.1 GI:3822114

KEYWORDS

```

SOURCE      Escherichia coli O157:H7
ORGANISM    Escherichia coli O157:H7
REFERENCE   1 (bases 1 to 92077)
AUTHORS     Blattner, F.R.
TITLE       The complete DNA sequence and analysis of the large virulence
JOURNAL     Plasmid of Escherichia coli O157:H7
MEDLINE     Nucleic Acids Res. 26 (18), 4196-4204 (1998)
PUBMED     98391744
REFERENCE   2 (bases 1 to 92077)
AUTHORS     Burland, V., Shao, Y., Perna, N.T., Plunkett, G. III, Sofia, H.J. and
            Blattner, F.R.
TITLE       Direct Submission
JOURNAL     Submitted (25-JUN-1998) Genetics, University of Wisconsin, 445
            Henry Mall, Madison, WI 53706, USA
FEATURES    Location/Qualifiers
             1..92077
               /organism="Escherichia coli O157:H7"
               /mol_type="genomic DNA"
               /strain="EDL933"
               /serotype="O157:H7"
               /db_xref="taxon:83334"
               /lab_host="Escherichia coli C600"
               /plasmid="pO157"
             1..561
               /genes="fno"
               /note="synonym: L7001"
             1..561
               /genes="fno"
               /note="97 pct identical amino acid sequence and equal
               length to fnoI_ECOLI SW: P22707"
               /codon_start=1
               /transl_table=11
               /product="fertility inhibition protein (conjugal transfer
               repressor)"
               /protein_id="AAC70069.1"
               /db_xref="GI:3822115"
               /translations="MAEQKRPVTLKRKTEGETPVRSKTIINVTTPPKWKKQKLA
               EKAREAEAAQAQALSIYINLPQDEAVNTLKPWPGFDGDTPLRLACGIRD
               VLEEDVAQRNIPLSHKLRRLAKAITRSSESYLCAMKAGACRYDTEGVVTEHISQEEEA
               YAAARDIKTRQRNRIKAEQLQALDEK"
             653..681
               /note="predicted sigma 70 promoter; score of 52%"
             698..949
               /genes="L7002"
             698..949
               /genes="L7002"
               /note="54 pct identical (1 gap) to 23 residues of an
               approx. 269 aa protein EXOK_RHIME SW: P33693"
               /codon_start=1
               /transl_table=11
               /product="unknown"
               /protein_id="AAC70070.1"
               /db_xref="GI:3822116"
               /translations="MSESTVHGTVRSVTSVPAGPLFWKSVDAQWKRQKHGDLGVL
               HPGLTGSSPEKGLNLTATGAEEGGNEKSLHYRDSQG"
             1025..1052
               /note="predicted sigma 70 promoter; score of 62%"
             1151..1612
               /genes="L7003"
             1151..1612
               /genes="L7003"
               /note="98 pct identical and equal length to yfi3_ECOLI SW:
               Q99342"
               /codon_start=1
               /transl_table=11
               /product="hypothetical protein 15.6 kDa protein in fno 3'
               region precursor"
               /protein_id="AAC70071.1"
               /db_xref="GI:3822117"

/translation="MRKIPLVLFIFSPVLCADIHGRVVRVLDGTIEVMDSRKAVR
IRLVNIDAPEKKQDYGRWSTDMMKSLVAGKTIVTVYFQDRYGRMLGVYVAPGMVNV
OPWRAGAWVYEQNTDPLVPLQNEARQQKRGWSDADPPPPWLRHRK"
1390..1419
/note="predicted sigma 70 promoter; score of 56%"
1658..1867
/gene="L7004"
1658..1867
/gene="L7004"
/note="55 pct identical (0 gaps) to 66 residues of an
approx. 72 aa protein HHA_ECOLI SW: P23870"
/codon_start=1
/transl_table=11
/product="putative hemolysin expression modulating
protein"
/protein_id="AAC70072.1"
/db_xref="GI:3822118"
/translation="MEKTKQEWLYQLRCCSVNTLEKIIHKNRDSLSTSERESFNAAA
DHRLAELITGKLYDRIPKEIKYVR"
1899..1927
/note="predicted sigma 70 promoter; score of 56%"
1905..2243
/gene="L7005"
1905..2243
/gene="L7005"
/note="98 pct identical (0 gaps) to 57 residues of an
approx. 200 aa protein; plasmid R100 miniplasmid pSM1 ORF
4, TRSW: Q52340"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAC70073.1"
/db_xref="GI:3822119"
/translation="MKLIIFILVLIALLIRILRSVNHSPLLMQLHAAGIRGTGD
AERILSGGBYDASHLRPERRRDILLEEVLQAGIPLLRSHDKRLQMTGEWINTTG
AAQSQSPHRS"
complement(2240..2327)
/note="initiation site; Escherichia coli plasmid R100 ssIB
gene; 99 pct identical to ssIB locus R10SSIB accession
D90185"
2416..2421
/note="100 pct identical (0 gaps) to the -35 region at
23..28 locus ECNR1RBP accession X02302"
2439..2445
/note="100 pct identical (0 gaps) to the -10 region at
46..52 locus ECNR1RBP accession X02302"
2473..2478
/note="100 pct identical (0 gaps) to RBS at 80..85 locus
ECNR1RBP accession X02302"
2483..2737
/gene="cpb2"
2483..2737
/note="synonym: L7006"
2483..2737
/gene="cpb2"
/note="98 pct identical amino acid sequence and equal
length to CPB2_ECOLI SW: P03847"
/codon_start=1
/transl_table=11
/product="CopB protein (RepA2 protein)"
/protein_id="AAC70074.1"
/db_xref="GI:3822120"
/translation="MSQTENAVTSSSGAKRAYRKGNPLSDAEKQRLSVARKRASFKEV
KVFLEPKYKAMLMQMCHEDGLTQAEVLTALIKSEAQRVCV"
2754..2782
/note="predicted sigma 70 promoter; score of 69%"
2973..3047
/gene="L7007"
2973..3047
/gene="L7007"
/note="91 pct identical (0 gaps) to 24 residues of an
approx. 24 aa protein uORF_F30REPFC, accession M16167,
translationally coupled to replication initiation protein"
/codon_start=1

```

```
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70075.1"
/db_xref="GI:3822121"
/transl_table="MLGKVQDFPLCSLLLRIVSAGWCD"
3032..3035
note="100 pct identical (0 gaps) to RBS at 641..644 locus
ECNR1REP accession X02302"
3040..3897
/gene="rep2"
/note="synonym: L7008"
3040..3897
/gene="rep2"
note="96 pct identical amino acid sequence and equal
length to REP2 ECOLI SW: P03066"
/codon_start=1
/transl_table=11
/product="replication initiation protein"
/protein_id="AAC70076.1"
/db_xref="GI:3822122"
/transl_table="MTDLQQTYYRQVKNPNVFPFREGAGTLKPKCKLMEKAVGFTSR
FDFAIHVAHARSKGLRRMPVLRRAIDALLQGLCFHYDPLANRVQCSITLAIKCG
LATSAAGKLSITRATLALFSLSELGLITYQTEYDPLIGCVIPTDIPTPALFAALDV
SEDAVAARSRVEMENLRKKQGLDITLGMDELIAKAWRFVREFRSYQTELKSRGFK
RARARDAGBERQDITLVKRLQLTREISEGRFTANREAVKREVRVVERMILSRNEN
YSRLATASP"
3847..4095
/standard_name="CIS"
/note="88 pct identical to locus ECCIS accession X12587,
required for cis-activation of oriR by the replication
initiation protein"
4072..4080
/note="dnaA site; 100 pct identical (0 gaps) to locus
ECNR1REP at (1682..1690) accession X02302"
4084..4232
/standard_name="oriR"
/note="89 pct identical to oriR (1094..1242); minimum
segment for replication of E. coli IncFII plasmid NRI
ECREPAL X12776"
/direction=right
4169..4197
/note="predicted sigma 70 promoter; score of 56%"
4259..4453
/gene="L7009"
4259..4453
/gene="L7009"
/note="95 pct identical to (0 gaps) 64 residues of a 128
aa protein REP4 locus ECRS1 accession V00351"
/codon_start=1

Query Match      86.2%; Score 1018; DB 1; Length 92077;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 11 TGGAAAAAGCCAAATTAATAATTCGCATCCAGCGCGCTCCAGCTGAAGTAGGCC 70
Db 19424 TGGAAAAAGCCAAATTAATAATTCGCATCCAGCGCGCTCCAGCTGAAGTAGGCC 19365

Qy 71 TGTTCCTGTCGGTATTTAAATGCAATGACCGTCCCGCTATTAAACAATGTGATAAATTA 130
Db 19364 TGTTCCTGTCGGTATTTAAATGCAATGACCGTCCCGCTATTAAACAATGTGATAAATTA 19305

Qy 131 CTCGGTTACCGGAAAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 190
Db 19304 CTCGGTTACCGGAAAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 19245

Qy 191 GTTCGATTTCCCTTAGCTGACTAGCCAGAGACACAATGATCTGTGCGGTTCTGTTAAT 250
Db 19244 GTTCGATTTCCCTTAGCTGACTAGCCAGAGACACAATGATCTGTGCGGTTCTGTTAAT 19185

Qy 251 ATCAACCGGTACTCAATATCTTCTCTGGGCTGGCTGCCATCATCCGGAAGCGTTCCGG 310
Db 19184 ATCAACCGGTACTCAATATCTTCTCTGGGCTGGCTGCCATCATCCGGAAGCGTTCCGG 19125
```

```
Qy 311 TCGGGATAAAAAATTCGCGCAGTGCAGCGGTCCTATGCAGACACATCCCCACGGGTAAACAG 370
Db 19124 TCGGGATAAAAAATTCGCGCAGTGCAGCGGTCCTATGCAGACACATCCCCACGGGTAAACAG 19065

Qy 371 CGTCCCTGTGCACATTTCTTCTGAATGACATCAGGAGTCCCGCCGTCCTCCTTGGCGATAAC 430
Db 19064 CGTCCCTGTGCACATTTCTTCTGAATGACATCAGGAGTCCCGCCGTCCTCCTTGGCGATAAC 19005

Qy 431 GGGCAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAACGCTTTCATTTCCGAGG 490
Db 19004 GGGCAGCGCGGAGACTGACGCTTTCAGCCAGTACCATACCAACGCTTTCATTTCCGAGG 18945

Qy 491 CATGACACACACTGCGCAATTCGCGTAGACCGGTAGACCGTGGGAAAAGGCGACCTGCCAT 550
Db 18944 CATGACACACACTGCGCAATTCGCGTAGACCGGTAGACCGTGGGAAAAGGCGACCTGCCAT 18885

Qy 551 TAAACATCTCCGCTCATATCCCAAGTGTCTGTCTGTCGACGACGAGCTGCTTCTGTAATTC 610
Db 18884 TAAACATCTCCGCTCATATCCCAAGTGTCTGTCTGTCGACGACGAGCTGCTTCTGTAATTC 18825

Qy 611 TTCACGCGCGCGCCCAACCAACGAGCCAGCGAAATGATTTCCCTTCCAATCTTCAGCTGATA 670
Db 18824 TTCACGCGCGCGCCCAACCAACGAGCCAGCGAAATGATTTCCCTTCCAATCTTCAGCTGATA 18765

Qy 671 CAATACACGCGAGCATAAATTCATGTCCTTTTCGGAGCGTAGCATCCCCACCTGGAACGAT 730
Db 18764 CAATACACGCGAGCATAAATTCATGTCCTTTTCGGAGCGTAGCATCCCCACCTGGAACGAT 18705

Qy 731 AAGCGGAACATTGCTGCTGATGCAGCCAGCGGTGATATGACGGGGTAAACGGTCCCAT 790
Db 18704 AAGCGGAACATTGCTGCTGATGCAGCCAGCGGTGATATGACGGGGTAAACGGTCCCAT 18645

Qy 791 GGCTTCATTATGCAATGCGGCGCCAGTCGAAACCGCGTGAATTAACCGTTACCGGTGCTCT 850
Db 18644 GGCTTCATTATGCAATGCGGCGCCAGTCGAAACCGCGTGAATTAACCGTTACCGGTGCTCT 18585

Qy 851 GACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGGCACAACAATGAATCACA 910
Db 18584 GACACCTTCGCGCATCAGATGCGCCATCATGCGGTGAGATAGGCACAACAATGAATCACA 18525

Qy 911 CAGATAAATTCAGGAAAAACGTTCTGCTTTCACGGGTGATGATGCTGTCTGACAAT 970
Db 18524 CAGATAAATTCAGGAAAAACGTTCTGCTTTCACGGGTGATGATGCTGTCTGACAAT 18465

Qy 971 AGTGAAGCGGTGACAGCATATCAGAGCGCTCAGTCCTGCTATATTAATGTCATGCCCACT 1030
Db 18464 ACTGAAGCGGTGACAGCATATCAGAGCGCTCAGTCCTGCTATATTAATGTCATGCCCACT 18405

Qy 1031 ATGGCAGATGACACAGATCAGGTTTAAATTCGCCGATTAATCCGTCGAAGTCTGAGGATGGA 1090
Db 18404 ATGGCAGATGACACAGATCAGGTTTAAATTCGCCGATTAATCCGTCGAAGTCTGAGGATGGA 18345

Qy 1091 AGGAAGGTGAAGCTGTCTCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTCTGGC 1150
Db 18344 AGGAAGGTGAAGCTGTCTCTGAAAGGAATAAAAGTGACATCATGCCCTCTTTTCTGGC 18285

Qy 1151 TTCGGAGCAATTTTACTTTTTTCTGCGAG 1181
Db 18284 TTCGGAGCAATTTTACTTTTTTCTGCGAG 18254
```

RESULT 6
AX191727/c

LOCUS AX191727 92077 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 9 from Patent WO0149775.
ACCESSION AX191727
VERSION AX191727.1 GI:15209896
KEYWORDS
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

```
REFERENCE
1
AUTHORS      Iversen, P.L.
TITLE        Antisense antibacterial cell division composition and method
JOURNAL      Patent: WO 0149775-A 9 12-JUN-2001;
              Avi Biopharma, Inc. (US)
FEATURES
  source     1. .92077
             /organism="Escherichia coli"
             /mol_type="unassigned DNA"
             /db_xref="taxon:562"
ORIGIN
Query Match      86.2%; Score 1018; DB 6; Length 92077;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 11 TGGAAAAAGCCAAATAAAAAATTGCCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 70
Db 19424 TGGAAAAAGCCAAATAAAAAATTGCCCATCCAGCGCGCTCCAGCTGAAAGTAGGCC 19365
Qy 71 TGTTCGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 130
Db 19364 TGTTCGTCCGGTATTTAAATGCAATGACCGTCCCGTATTTAAACAATGTGATAATTA 19305
Qy 131 CTCGGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCCGGTTATCT 190
Db 19304 CTCGGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCCGGTTATCT 19245
Qy 191 GTTGCAATTCCTTAGCTGACTAGCCAGAGACACAATGATCTGTGCGCTTCTGTGTAAT 250
Db 19244 GTTGCAATTCCTTAGCTGACTAGCCAGAGACACAATGATCTGTGCGCTTCTGTGTAAT 19185
Qy 251 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 310
Db 19184 ATCAAAACCGGTACTCAATATCTTCTGCGCTGGCTGCCATCATCCGGAAGCGTTCCGG 19125
Qy 311 TCGGGATAAAAAATTCGCGAGTGGCGGTCCTATGACAGACACATCCCCACGGGTAAACAG 370
Db 19124 TCGGGATAAAAAATTCGCGAGTGGCGGTCCTATGACAGACACATCCCCACGGGTAAACAG 19065
Qy 371 CGTCCCTGTACATCTTCTGTAATGACATCAGGGATCCCGCGTCTCACTGGCGATAAC 430
Db 19064 CGTCCCTGTACATCTTCTGTAATGACATCAGGGATCCCGCGTCTCACTGGCGATAAC 19005
Qy 431 GGGCAGCCCGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCATTTTCCGAAGG 490
Db 19004 GGGCAGCCCGAGACTGACGCTTCAGCCAGTACCATAACAAACGCTTCATTTTCCGAAGG 18945
Qy 491 CATGACCAACACACTGGCAATCCGGTAGACCGGTAAACGCTGGGAAAAGGCACTGCGCAT 550
Db 18944 CATGACCAACACACTGGCAATCCGGTAGACCGGTAAACGCTGGGAAAAGGCACTGCGCAT 18885
Qy 551 TAACACATCTCCGTCATTTCCAGGTGTTCTGTCTGTGACGACGAGCTGCTTCGTATTC 610
Db 18884 TAACACATCTCCGTCATTTCCAGGTGTTCTGTCTGTGACGACGAGCTGCTTCGTATTC 18825
Qy 611 TTACGCGCCGCGGCCACCAACGAGCAGCAATGATTTCCCTTCCATCTTCAGCTGATA 670
Db 18824 TTACGCGCCGCGGCCACCAACGAGCAGCAATGATTTCCCTTCCATCTTCAGCTGATA 18765
Qy 671 CAATACACGAGCATAAATTCATGTCTCTTTTCGGGACGTAGCATCCCACTGAAACGAT 730
Db 18764 CAATACACGAGCATAAATTCATGTCTCTTTTCGGGACGTAGCATCCCACTGAAACGAT 18705
Qy 731 AAGCGGAACATTTCTGCTGATGACGCGCCAGGCTGGATATGACGGGTAAACGGTCCGAT 790
Db 18704 AAGCGGAACATTTCTGCTGATGACGCGCCAGGCTGGATATGACGGGTAAACGGTCCGAT 18645
Qy 791 GGCTTCATTATGCAATCGCGGCCAGTCCGAAACCCGGTGGATAACCGGTATACCGGTGTCCT 850
Db 18644 GGCTTCATTATGCAATCGCGGCCAGTCCGAAACCCGGTGGATAACCGGTATACCGGTGTCCT 18585
Qy 851 GACACCTTCCGCCATCAGATGCGGCCATCATATGGGTGAGATAGGCACAAACAATGAAATCACA 910
```

```
Db 18584 GACACCTTCCGCCATCAGATGCGCCATCATGGCTAGGCACCAACAATGAAATCACA 18525
Qy 911 CAGATAATTCAGGAAACGTTCTGCTCTACGGGTGATGCTAGGTTTGTCTGACAAT 970
Db 18524 CAGATAATTCAGGAAACGTTCTGCTCTACGGGTGATGCTAGGTTTGTCTGACAAT 18465
Qy 971 AGTGAACGGGTGACAGCATATCAGACGGGTCACTCTGCTATATTAATCTGTCTGGCCACT 1030
Db 18464 ACTGAACGGGTGACAGCATATCAGACGGGTCACTCTGCTATATTAATCTGTCTGGCCACT 18405
Qy 1031 ATGGCAGATGACAGATCAGGTTTAAATTCGCCGATAATCCCGTCAAGTCTGAGGATGGA 1090
Db 18404 ATGGCAGATGACAGATCAGGTTTAAATTCGCCGATAATCCCGTCAAGTCTGAGGATGGA 18345
Qy 1091 AGGAAGGTGAGCGCTTCTCTGAAAGGAATAAAGTGACATCATGCCCTCTTTTCTGGC 1150
Db 18344 AGGAAGGTGAGCGCTTCTCTGAAAGGAATAAAGTGACATCATGCCCTCTTTTCTGGC 18285
Qy 1151 TTCGGGAGCAATTTTACTTTTCTCTGCAG 1181
Db 18284 TTCGGGAGCAATTTTACTTTTCTCTGCAG 18254
RESULT 7
AB011549/c
LOCUS      AB011549
DEFINITION Escherichia coli plasmid pO157 DNA, complete sequence.
ACCESSION AB011549
VERSION    AB011549.2 GI:4589740
KEYWORDS   ToxR-regulated lipoprotein; tagA.
SOURCE     Escherichia coli
ORGANISM   Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.
REFERENCE  1 (sites)
AUTHORS    Makino, K., Ishii, K., Yasunaga, T., Hattori, M., Yokoyama, K.,
            Yutsudo, H.C., Kubota, Y., Yamaichi, Y., Iida, T., Yamamoto, K.,
            Honda, T., Han, C.G., Ohtsubo, E., Kasamatsu, M., Hayashi, T., Kuhara, S.
            and Shinagawa, H.
TITLE      Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an
            enterohemorrhagic Escherichia coli O157:H7 derived from Sakai
            outbreak
JOURNAL    DNA Res. 5 (1), 1-9 (1998)
MEDLINE    98290540
PUBMED     9628576
REFERENCE  2 (bases 1 to 92721)
AUTHORS    Makino, K.
TITLE      Direct Submission
JOURNAL    Submitted (24-FEB-1998) Kozo Makino, Research Institute for
            Microbial Diseases, Osaka University, Molecular Microbiology;
            Yamadaoka, 3-1, Suita, Osaka 562, Japan
            E-mail:makinobk@biken.osaka-u.ac.jp, Tel:81-6-879-8318,
            Fax:81-6-879-8320)
COMMENT    On Apr 20, 1999 this sequence version replaced gi:3336997.
FEATURES
  source     1. .92721
             /organism="Escherichia coli"
             /mol_type="genomic DNA"
             /strain="O157:H7"
             /sub_strain="RIMD 0509952"
             /db_xref="taxon:562"
             /plasmid="pO157"
             /notes="RIMD 0509952 is a strain of enterohemorrhagic E.
            coli, EHEC O157:H7"
            join(92527..92721,1..2502)
            /genes="tagA"
            join(92527..92721,1..2502)
            /genes="tagA"
            /codon_start=1
            /transl_table=11
            /product="ToxR-regulated lipoprotein"
            /protein_id="BAA31757.3"
  gene
  CDS
```

```
/db_xref="GI:4666293"
/translation="MNTKQNRWRTPMKLIKLSCTTILAPLAIGVFSATAADNNSAIYF
NTSQINDLQGSAAEVFAQSQILPAHPKSGDSQPHLTSLRSKSLLLVVRPADDKTP
VQVEARDNDGILGTLTYPPSLPTDIYHLDGPEGGIDFTPHNGTKIKIINTVAEVN
KLSDAGSSISHSLTNNALVEIHTANGRWVRDIYLPQGPDLGKMRVFSYSSSTV
YGDVKVTLVSGVNTLPKPYNGQWFRSGLENNRIYTAQHIWSAELPAHWIYPLNLV
IKQGNLSGRNDIKIGAPGELLHTIDIGLTT'PRORFPAKDKAHEVFPQIPVSR
MIYNYAPLHKELVMTPTGELLTMDPGNGGWSGTWRQRIKELVSHGIDNNANYGLN
STAGLGNSHPYVQAUAHNSRGNTANGIQVHGSGGGIYVLDSTLGNBFSHEVGH
NYGLHYVDKFGSVKHSVAENNSNNTWGDGDKRFIPNFYPSQTNKSKLNNOCQFP
DGHKFGFDAGSPFSAANRFTWYTPNSSAIQRFPENKAVFDSRSSTGYKSNADT
QMEPEYHTIDRAEQITASVNELSEKMAELMAEYAVVKVMNGNWTNRIYIPTASA
QNSGILITINHEAGYSYLFINGDEKVVSGYKKS'FVSDQCFKEDVDVDTREARKE
QFGVPVTLVGYDPEGTLSSYIPAMYGIFTYSDDSNLSDNDCLQGVQEQQL
RFLNHRANNTVNMKPHINVPSTEQATLVNKNKILDTKSLTAPAPGLTYTVNGQ
ALPAKNEGCVIVNSGKCYCLFVQQRSGYSLPDWVGQEVVYDVGAKAKVLLSDWDN
LSYNRICEFVGNVPADMKKVKAWNGQYLD'FSKPRSMRVYK"
gene
2589. .3464
CDS
/gene="etpC"
2589. .3464
/gene="etpC"
/codon_start=1
/translation="MNFUSFRGDRGLFIKDIVLKM'LTNRLLCVILLIAGYQLVSVI
HFPLTQASVPGLSRVSAPETAVTGDQTEERFVTLFGRASPLSSEGRAQETMPSLS
DILLSEGDLDVRLGILYSVAEHSVAIPAHNNRQFSLSGVEKVPYDATISAI'PSDHI
INYGKTVSLPRYDNTKKNAYDNNLTVGDVITQDNFRVESVFDIMSFSAVTNNT
LSGYRLIPGHSSLYFNAGLHDNDLAVSYNGSLRTRQQAQIMKQLPELKEIKITVE
RDQLDYDAFVGEN"
gene
3675. .5432
CDS
/gene="etpD"
3675. .5432
/gene="etpD"
/codon_start=1
/translation="MLNEEQYQFFLSVLDVYGVAVVDMHNGILKVVRSKDAKTSVAP
VASDVSQGVDEVTVVVPVSVVAERDLAPLLQDNDAGASVHVYERSVNLLMTGR
AAYMKRLMEIVERDKVGRSVATP'LTASATDVARLVLTETKTETDKTAIPAMTAK
LVADERTNVLSGEPISQORIIISIKQLDRQDVQGNKVIYLYKAKOLDVILVTG
ISSIENDSKSPSTEARLKGVIKISHEQTNALILTCAPDIRLNVISQLDIRPQ
VLVEAIIAETQADAGNLGIQWYNHAGVAQFTSTGLPIT'WVQTRQNEILSDSDNSA
LSMENGIAAGFYQGNWMLLTALSTSSKNDILATPSIVLTDNNMEATFNVOEVPVLSG
SQTSGDNI'FNTVERTVGIKLRVQKINEGDSVLEIEQVSGVADTAVATTDLGA
TFNTRVTNMLVNGNETVVVGGLLDKSINGSESKVPLIGDIPVIGHLFRKSEQTKA
RNLMLFIRPTIIRERDGRFRAHSAEKYQSFNQEQVQSRGKETTALT'LNERQLRLSPDQ
DTAFRKVAALAFYAQEM"
gene
5432. .6937
CDS
/gene="etpE"
5432. .6937
/gene="etpE"
/codon_start=1
/translation="MSRQVNVSESRPLLPFSR'QRIILLREQNGNRVFCMEDTPA
SALLEVRRVAEGPLNVTYSAAEFQKLVSSYQSDDEARQMAE'IGNEMDFYVAGE
LPREDLLDANDDAP'IIRLINAMLEAIKEKASDIHETV'ERHLVRFPIDGVLREIL
RLHRNLASILIRIKYMAKRLDIAEKRPQDGRMLRIGRAVDVRVSTVSNHGERIV
LRLLDKNSVSLDLAALGMSQONORHIDALIRPHGII'LVTPGTGSKSTLYAALSLL
NPRDNIMVEDPVEV'ELGISQTVNPKVMT'FARSLAIRLRQDPDVVLYGEIRDGE
TAQIAVASLTGLHLVSTLHNSAAGLSRLQMDGIP'PELLSTSLAVLAQRLVTLIC
PRCQPCQVSTELAMDMDIP'PETTIWQAGCQCHSFTGYHRTGHELLIDDRIRTA
IYQEGELGITRLAGSRYLTLRGDGRQKVLAGE'TSWEVVRVVTESLQEEH"
gene
6939. .8162
CDS
/gene="etpF"
6939. .8162
```

```
/gene="etpF"
/codon_start=1
/translation="MNFUSFRGDRGLFIKDIVLKM'LTNRLLCVILLIAGYQLVSVI
HFPLTQASVPGLSRVSAPETAVTGDQTEERFVTLFGRASPLSSEGRAQETMPSLS
DILLSEGDLDVRLGILYSVAEHSVAIPAHNNRQFSLSGVEKVPYDATISAI'PSDHI
INYGKTVSLPRYDNTKKNAYDNNLTVGDVITQDNFRVESVFDIMSFSAVTNNT
LSGYRLIPGHSSLYFNAGLHDNDLAVSYNGSLRTRQQAQIMKQLPELKEIKITVE
RDQLDYDAFVGEN"
gene
3675. .5432
CDS
/gene="etpD"
3675. .5432
/gene="etpD"
/codon_start=1
/translation="MLNEEQYQFFLSVLDVYGVAVVDMHNGILKVVRSKDAKTSVAP
VASDVSQGVDEVTVVVPVSVVAERDLAPLLQDNDAGASVHVYERSVNLLMTGR
AAYMKRLMEIVERDKVGRSVATP'LTASATDVARLVLTETKTETDKTAIPAMTAK
LVADERTNVLSGEPISQORIIISIKQLDRQDVQGNKVIYLYKAKOLDVILVTG
ISSIENDSKSPSTEARLKGVIKISHEQTNALILTCAPDIRLNVISQLDIRPQ
VLVEAIIAETQADAGNLGIQWYNHAGVAQFTSTGLPIT'WVQTRQNEILSDSDNSA
LSMENGIAAGFYQGNWMLLTALSTSSKNDILATPSIVLTDNNMEATFNVOEVPVLSG
SQTSGDNI'FNTVERTVGIKLRVQKINEGDSVLEIEQVSGVADTAVATTDLGA
TFNTRVTNMLVNGNETVVVGGLLDKSINGSESKVPLIGDIPVIGHLFRKSEQTKA
RNLMLFIRPTIIRERDGRFRAHSAEKYQSFNQEQVQSRGKETTALT'LNERQLRLSPDQ
DTAFRKVAALAFYAQEM"
gene
5432. .6937
CDS
/gene="etpE"
5432. .6937
/gene="etpE"
/codon_start=1
/translation="MSRQVNVSESRPLLPFSR'QRIILLREQNGNRVFCMEDTPA
SALLEVRRVAEGPLNVTYSAAEFQKLVSSYQSDDEARQMAE'IGNEMDFYVAGE
LPREDLLDANDDAP'IIRLINAMLEAIKEKASDIHETV'ERHLVRFPIDGVLREIL
RLHRNLASILIRIKYMAKRLDIAEKRPQDGRMLRIGRAVDVRVSTVSNHGERIV
LRLLDKNSVSLDLAALGMSQONORHIDALIRPHGII'LVTPGTGSKSTLYAALSLL
NPRDNIMVEDPVEV'ELGISQTVNPKVMT'FARSLAIRLRQDPDVVLYGEIRDGE
TAQIAVASLTGLHLVSTLHNSAAGLSRLQMDGIP'PELLSTSLAVLAQRLVTLIC
PRCQPCQVSTELAMDMDIP'PETTIWQAGCQCHSFTGYHRTGHELLIDDRIRTA
IYQEGELGITRLAGSRYLTLRGDGRQKVLAGE'TSWEVVRVVTESLQEEH"
gene
6939. .8162
CDS
/gene="etpF"
6939. .8162
```

/translation="MKLREQVALLVLLILSLMVTIAAVIAERNGRFTLRVAQLDQ
LOAKWGYTAETIAKOLIORSESPRKTHLAONWASOROPETRGDVRGOIVDAQA
CENLAINVYGVDLTISIPYAARIFQOLLINLOVELLOARQVTFALRDMIDRDDKPVKG
GADEVYMGPEPFLAANQWQDVSELRILRGIDARKYKLLPYVCVLPFSDLSNVVN
TLDSQAPLLAALFLTKPDLSPVTELLQRRPRTGWSVAAPLADAPLPKIDIDTSAAMPV
LAVSSNYFLVRLHVRSGEHLFSQOTLMQWREERFRIRIQYGLTMBEV"
11473. .12321

gene

Query Match 86.2%; Score 1018; DB 1; Length 92721;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 70
DB 88939 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 88880

QY 71 TGTCTGTCCGGTATTTAAATGCAATGACCGTCCCGGTATTTAAACAATGTGATAATTA 130
DB 88879 TGTCTGTCCGGTATTTAAATGCAATGACCGTCCCGGTATTTAAACAATGTGATAATTA 88820

QY 131 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 190
DB 88819 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 88760

QY 191 GTTGCAATTTCCCTTAGCCCTGACTAGCGAGACACAATGATCTGTGCCGTTCTGTGTAAT 250
DB 88759 GTTGCAATTTCCCTTAGCCCTGACTAGCGAGACACAATGATCTGTGCCGTTCTGTGTAAT 88700

QY 251 ATCAACCGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCTCGGAAGCGTTCCGG 310
DB 88699 ATCAACCGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCTCGGAAGCGTTCCGG 88640

QY 311 TCCGGATAAAAAATCCGCGAGTCCGCGTCCATGACAGACACATCCCGACGAGTAACAG 370
DB 88639 TCCGGATAAAAAATCCGCGAGTCCGCGTCCATGACAGACACATCCCGACGAGTAACAG 88580

QY 371 CGTCCCTGTGCACATCTTCTTGAATGACATCAGGATCCCGCGCTGTCTCACTGGCGATAAC 430
DB 88579 CGTCCCTGTGCACATCTTCTTGAATGACATCAGGATCCCGCGCTGTCTCACTGGCGATAAC 88520

QY 431 GGGCAGCCGGAGACTGACGCTTCAGCAGTACCATACCAAAACGCTTCATTTCCGAAGG 490
DB 88519 GGGCAGCCGGAGACTGACGCTTCAGCAGTACCATACCAAAACGCTTCATTTCCGAAGG 88460

QY 491 CATGACCACACACTGCAATCCGCTAGACCGGTAGCGTGAAGGAGGACCTGCGCAT 550
DB 88459 CATGACCACACACTGCAATCCGCTAGACCGGTAGCGTGAAGGAGGACCTGCGCAT 88400

QY 551 TAACACATCTCCGTCATATCCGAGGTCTTCTGTCTGTGACGAGAGCTGCTTCGTATTC 610
DB 88399 TAACACATCTCCGTCATATCCGAGGTCTTCTGTCTGTGACGAGAGCTGCTTCGTATTC 88340

QY 611 TTACGCGCGCGGCCACACGAGCGAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 670
DB 88339 TTACGCGCGCGGCCACACGAGCGAGCGAATGATTTCCCTTCCATCTTCAGCTGATA 88280

QY 671 CAATACACGAGCATAAATTCATGCTCTTTTTCGGGACGTAGCATCCCGACCTGAAACGAT 730
DB 88279 CAATACACGAGCATAAATTCATGCTCTTTTTCGGGACGTAGCATCCCGACCTGAAACGAT 88220

QY 731 AAGCGGAACATTTGTCTGATGACGCCCGAGGCGTGGATATGACGGGGTAAACCGTGCAT 790
DB 88219 AAGCGGAACATTTGTCTGATGACGCCCGAGGCGTGGATATGACGGGGTAAACCGTGCAT 88160

QY 791 GGCTTCATTTATGCAATCGGGCCAGTCCGAAACCGCGTGGATATACCGTTACCGGTGTCCT 850
DB 88159 GGCTTCATTTATGCAATCGGGCCAGTCCGAAACCGCGTGGATATACCGTTACCGGTGTCCT 88100

QY 851 GACACCTTCGCGCATCAGATCGGCATCATGCGGTGAGATAGGCACACAATGAAATCACA 910
DB 88099 GACACCTTCGCGCATCAGATCGGCATCATGCGGTGAGATAGGCACACAATGAAATCACA 88040

QY 911 CAGATAATTGAGGAAAAACGTTCTGGTCTTACGCGGTGATGTAGGTTTTTTGTCTGACAAAT 970

DB 88039 CAGATAATTTCAGGAAAAAGCTTCTGTCTTACGGGTGATGTAGTTTTTTGTCTGACAAAT 87980

QY 971 AGTGAAGCCGTGACAGCATATCAGAGCGCTCAGTCCGCTGCTATATTAATCTGTCATGGCCACT 1030

DB 87979 ACTGAAGCCGTGACAGCATATCAGAGCGCTCAGTCTGCTATATTAATCTGTCATGGCCACT 87920

QY 1031 ATGGCAGATGACCCAGATCAGGTTTAAATTCGCCGATAATCCCGTGAAGTCTGAGGATGGA 1090

DB 87919 ATGGCAGATGACCCAGATCAGGTTTAAATTCGCCGATAATCCCGTGAAGTCTGAGGATGGA 87860

QY 1091 AGGAAGGTGAAGCTCTTCTCTGAAGGAATAAAGTGAACATCATGCGCTCTTTTCTGGC 1150

DB 87859 AGGAAGGTGAAGCTCTTCTCTGAAGGAATAAAGTGAACATCATGCGCTCTTTTCTGGC 87800

QY 1151 TTCCGAGCAATTTTACTTTTCTCTGCGAG 1181

DB 87799 TTCCGAGCAATTTTACTTTTTTCTCTGCGAG 87769

RESULT 8
AX191725/c
LOCUS AX191725 92721 bp DNA linear PAT 15-AUG-2001
DEFINITION Sequence 7 from Patent WO0149775.
ACCESSION AX191725
VERSION AX191725.1 GI:15209894
KEYWORDS Escherichia coli
SOURCE Escherichia coli
ORGANISM Escherichia coli
Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
Enterobacteriaceae; Escherichia.

REFERENCE
1
Iversen, P.L.
Antisense antibacterial cell division composition and method
Patent: WO 0149775-A 7 12-JUL-2001;
Avi Biopharma, Inc. (US)
Location/Qualifiers
1. .92721
/organism="Escherichia coli"
/mol_type="unassigned DNA"
/db_xref="taxon:562"

ORIGIN
Query Match 86.2%; Score 1018; DB 6; Length 92721;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1168; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 70
DB 88939 TGGAAAAAGCCAAATAAAAAATTCGCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 88880

QY 71 TGTCTGTCCGGTATTTAAATGCAATGACCGTCCCGGTATTTAAACAATGTGATAATTA 130
DB 88879 TGTCTGTCCGGTATTTAAATGCAATGACCGTCCCGGTATTTAAACAATGTGATAATTA 88820

QY 131 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 190
DB 88819 CTCGGTTACCGGAAACCGCTGAACAAAATTCGGGCTGAAAGAGGATCCGCGTTATCT 88760

QY 191 GTTGCAATTTCCCTTAGCCCTGACTAGCGAGACACAATGATCTGTGCCGTTCTGTGTAAT 250
DB 88759 GTTGCAATTTCCCTTAGCCCTGACTAGCGAGACACAATGATCTGTGCCGTTCTGTGTAAT 88700

QY 251 ATCAACCGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCTCGGAAGCGTTCCGG 310
DB 88699 ATCAACCGGTACTCAATATCTTCTGTGGCGTGGCTGCCATCATCTCGGAAGCGTTCCGG 88640

QY 311 TCCGGATAAAAAATCCGCGAGTCCGCGTCCATGACAGACACATCCCGACGAGTAACAG 370
DB 88639 TCCGGATAAAAAATCCGCGAGTCCGCGTCCATGACAGACACATCCCGACGAGTAACAG 88580

QY 371 CGTCCCTGTGCACATCTTCTTGAATGACATCAGGATCCCGCGCTGTCTCACTGGCGATAAC 430

/product="unknown"
/protein_id="AAC24351.1"
/db_xref="GI:3253292"
/translation="MLNARHLPLVLMYHHVSRCPGLVTLSPETFEQMKWLADNNWRITV
TSABEYFYGGTLPKRSVMLTDFDGYLDNWLRAWPLVKEYNLHAHIFLITGLIGKN
VRSRQOEYSHRDCERLIADNRSDENVLRWSEVRMDSLGVEFHVHTSHKRWDLKS
VSRAEQRLMKEDILVKGQCLTKELGFCFSHLCPWEGYNNRDYINLAGKLGSYLVTY
ERRMNCPEGSLRIGRISTKEREHSGLKRLRFLYYTTPLFSSVLALHKGPRLPDN"

ORIGIN

Query Match 77.6%; Score 916; DB 1; Length 5612;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1166; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 11 TGGAAAAAGCCAAATAAATAATGCCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 70
DB 2415 TGGAAAAAGCCAAATAAATAATGCCCATCCAGCGCGCTCCAGCTGAAGTAGGCC 2474

QY 71 TGTTCCTGTCGGGTATTTAAATGCATTGACCGTCCCGGTATTTAAACAATGTGATAATTA 130
DB 2475 TGTTCCTGTCGGGTATTTAAATGCATTGACCGTCCCGGTATTTAAACAATGTGATAATTA 2534

QY 131 CTCGGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCGTTATCT 190
DB 2535 CTCGGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAGAGGATCCGCGTTATCT 2594

QY 191 GTTGCAATTTCCCTTAGCCTGAGTCCAGCAGACACAATGATCTGTGCGGTTCTGTGTAAT 250
DB 2595 GTTGCAATTTCCCTTAGCCTGAGTCCAGCAGACACAATGATCTGTGCGGTTCTGTGTAAT 2654

QY 251 ATCAAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCGGAAGCGTTCCGG 310
DB 2655 ATCAAAACCGGTACTCAATATCTTCTCTGGCGCTGGCTGCCATCATCGGAAGCGTTCCGG 2714

QY 311 TCGGGATAAAAATCGCGAGTGGCGGCTGCATGACAGACACATATCCGCCACCGGGTAACAG 370
DB 2715 TCGGGATAAAAATCGCGAGTGGCGGCTGCATGACAGACACATATCCGCCACCGGGTAACAG 2774

QY 371 CGTCCCTGTACATCTTCTGATGATGACATCAGAGATCCCGCGCTCTCACTGCGGATAAC 430
DB 2775 CGTCCCTGTACATCTTCTGATGATGACATCAGAGATCCCGCGCTCTCACTGCGGATAAC 2834

QY 431 GGGCAGCGCGGAGACTGACGCTTCAGCCAGTACCACACAAAGCGTTCAITTTCCGAAGG 490
DB 2835 GGGCAGCGCGGAGACTGACGCTTCAGCCAGTACCACACAAAGCGTTCAITTTCCGAAGG 2894

QY 491 CATGACACACACTGGCAATCCGGTAGACCGGTAGACCGGTGGGAAAAGGCACTGCGCAT 550
DB 2895 CATGACACACACTGGCAATCCGGTAGACCGGTAGACCGGTGGGAAAAGGCACTGCGCAT 2954

QY 551 TAACACATCTCCGTCATTTCCAGGTGTTCTGTCTGTGACGACGAGTGTCTGCTATTC 610
DB 2955 TAACACATCTCCGTCATTTCCAGGTGTTCTGTCTGTGACGACGAGTGTCTGCTATTC 3014

QY 611 TTCACGCCCGCGGCCACACAGCCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA 670
DB 3015 TTCACGCCCGCGGCCACACAGCCAGCGAAATGATTTCCCTTCCATCTTCAGCTGATA 3074

QY 671 CAATACACGAGCATAAATTCATGTCTTTTTCGGACGTAGACATCCCACTGAACGAT 730
DB 3075 CAATACACGAGCATAAATTCATGTCTTTTTCGGACGTAGACATCCCACTGAACGAT 3134

QY 731 AAGCGGAACATTGTCTGCTGATCAGCCAGCGGTGGATATGACAGGGTAACCGTCGCAT 790
DB 3135 AAGCGGAACATTGTCTGCTGATCAGCCAGCGGTGGATATGACAGGGTAACCGTCGCAT 3194

QY 791 GGCTTCATTAATGCAATGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGTCT 850
DB 3195 GGCTTCATTAATGCAATGCGGCCAGTCGAAACCCGGTGGAAATACCGTTACCGGTGTCT 3254

QY 851 GACACCTTCGCCCATCGATCGGCCATCATGGGTGAGATAGGCACAACAATGAATCACA 910
DB 3255 GACACCTTCGCCCATCGATCGGCCATCATGGGTGAGATAGGCACAACAATGAATCACA 3314

QY 911 CAGATAATTTCAGGAAAAAGCTTCTGGTCTTACGGGTGATGTAGGTTTTTTTGTCTGACAAT 970
DB 3315 CAGATAATTTCAGGAAAAAGCTTCTGGTCTTACGGGTGATGTAGGTTTTTTTGTCTGACAAT 3374

QY 971 AGTGAAGCGGTGACACATATCAGACGGCTCAGTCTGCTATATTAATCTGTCATGGCCACT 1030
DB 3375 ACTGAAGCGGTGACACATATCAGACGGCTCAGTCTGCTATATTAATCTGTCATGGCCACT 3434

QY 1031 ATGGCAGATGACACAGATCAGGTTTAAATTTCCCGATAATCCGTCGAAGTCTGAGGATGA 1090
DB 3435 ATGGCAGATGACACAGATCAGGTTTAAATTTCCCGATAATCCGTCGAAGTCTGAGGATGA 3494

QY 1091 AGGAAGGTGAAGCTGTCTCTGAAAGCAATAAAAGTGACATCATGCGCTCTTTTCTGGC 1150
DB 3495 AGGAAGGTGAAGCTGTCTCTGAAAGCAATAAAAGTGACATCATGCGCTCTTTTCTGGC 3554

QY 1151 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 1181
DB 3555 TTCCGGAGCAATTTTACTTTTTTCTCTGCAG 3585

RESULT 10
BD205235
LOCUS
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC).
ACCESSION BD205235
VERSION BD205235.1 GI:33015005
KEYWORDS JP 2002512813-A/25.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 31)
AUTHORS Frechon,D.T.M., Laure,F.C. and Thierry,D.
TITLE Nucleotide sequence for detecting enterohemorrhagic Escherichia coli (EHEC)
JOURNAL Patent: JP 2002512813-A 25 08-MAY-2002;
COMMENT BIORAD PASTEUR
OS Unidentified
PN JP 2002512813-A/25
PD 08-MAY-2002
PR 27-APR-1999 JP 2000546051
PR 28-APR-1998 FR 98/05329
PI DOMINIQUE THERESE MARIE FRECHON, FRANCOISE CLAUDINE LAURE, PI
DOMINIQUE THIERRY
PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/69,C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
CC Nucleotide sequence for detecting enterohemorrhagic CC
Escherichia coli
CC (EHEC).
FH Key Location/Qualifiers
FT source 1..31
FT /organism='Unidentified'.
FEATURES
source
1..31
/organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 2.6%; Score 31; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 928 ACCTTCTGGTCTTACGGGTGATGTAGGTTTT 958
DB 1 ACCTTCTGGTCTTACGGGTGATGTAGGTTTT 31

RESULT 11
BD205236


```
LOCUS      BD205236                      31 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Nucleotide sequence for detecting enterohemorrhagic Escherichia
            coli (EHEC).
ACCESSION  BD205236
VERSION    BD205236.1 GI:33015006
KEYWORDS   JP 2002512813-A/26.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 31)
AUTHORS    Frechon,D.T.M., Laure,F.C. and Thierry,D.
TITLE      Nucleotide sequence for detecting enterohemorrhagic Escherichia
            coli (EHEC)
JOURNAL    Patent: JP 2002512813-A 26 08-MAY-2002;
COMMENT    BIORAD PASTEUR
           OS Unidentified
           PN JP 2002512813-A/26
           PD 08-MAY-2002
           PF 27-APR-1999 JP 2000546051
           PR 28-APR-1998 FR 98/05329
           PI DOMINIQUE THERESE MARIE FRECHON,FRANCOISE CLAUDE LAURE, PI
           PC C12N9/08,C07K14/245,C12N1/21,C12N15/09,C12Q1/68,C12N15/00 CC
           Strandedness: Single;
           CC Topology: Linear;
           CC Nucleotide sequence for detecting enterohemorrhagic CC
           CC Escherichia coli
           PH Key
           FT source
           FT Location/Qualifiers
FEATURES   source
            1..31
            Location/Qualifiers
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
Query Match      2.6%; Score 31; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      970 TAGTGAAGCGGTGACAGCATATCAGACGGCT 1000
       |||||||
Db      1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31

RESULT 12
AX011321
LOCUS      AX011321                      31 bp      DNA      linear      PAT 06-SEP-2000
DEFINITION Sequence 25 from Patent WO955908.
ACCESSION  AX011321
VERSION    AX011321.1 GI:9997871
KEYWORDS   Escherichia coli
SOURCE     Escherichia coli
ORGANISM   Escherichia coli
REFERENCE  1
AUTHORS    Thierry,D., Frechon,D.T. and Laure,P.C.
TITLE      Nucleotide sequences for detecting enterohemorrhagic escherichia
            coli (ehec)
JOURNAL    Patent: WO 955908-A 25 04-NOV-1999;
           THIERRY DOMINIQUE (FR); FRECHON DOMINIQUE THERESE MARI (FR); LAURE
           FRANCOISE CLAUDE (FR); PASTEUR SANOFI DIAGNOSTICS (FR)
FEATURES   source
            1..31
            Location/Qualifiers
            /organism="Escherichia coli"
            /mol_type="unassigned DNA"
            /db_xref="taxon:562"
ORIGIN
Query Match      2.6%; Score 31; DB 6; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      970 TAGTGAAGCGGTGACAGCATATCAGACGGCT 1000
       |||||||
Db      1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31

RESULT 14
BX571861/c
LOCUS      BX571861/c                    342905 bp    DNA      linear      BCT 26-SEP-2003
DEFINITION Photorhabdus luminescens subsp. laumondii TT01 complete genome;
            segment 3/17.
ACCESSION  BX571861 BX470251
VERSION    BX571861.1 GI:36784041
KEYWORDS   complete genome.
SOURCE     Photorhabdus luminescens subsp. laumondii TT01
ORGANISM   Photorhabdus luminescens subsp. laumondii TT01
REFERENCE  1
AUTHORS    Duchaud,E., Rusniok,C., Frangeul,L., Buchrieser,C., Taourit,S.,
           Bocs,S., Bouraux-Eude,C., Chandrai,M., Dassa,E., Derose,R.,
           Derzelle,S., Freysinet,G., Gaudriault,S., Givaudan,A., Glaser,P.,
           Medigue,C., Lanois,A., Powell,K., Siquier,P., Wingate,V.,
           Zouine,M., Boemare,N., Danchin,A. and Kunst,P.
           Complete genome sequence of the entomopathogenic bacterium
           Photorhabdus luminescens
           Nat. Biotechnol. 11 (1) (2003) In press
           2
           Duchaud,E., Frangeul,L., Rusniok,C. and Kunst,F.
           Direct Submission
           Submitted (23-APR-2003) L. Frangeul, Institut Pasteur, Genopole, 25
           rue du Docteur Roux, 75724 Paris Cedex 15, FRANCE. E-mail:
           lfrangeu@pasteur.fr, fkunst@pasteur.fr
           Location/Qualifiers
           1..342905
           /organism="Photorhabdus luminescens subsp. laumondii TT01"
           /mol_type="genomic DNA"
           source
```


Search completed: June 4, 2005, 17:17:23
Job time : 5486 secs

VFTEPEPLISQEVYKQIIEBKALSRIAEGDVEQEARLKSGERILVOLNAGLSLKVELQ
ISNSIDGVGLYRTEIFPMFLGCFPEDEQKARYOEMQLQFPDPKPVVLRITLDIGADKQL
PYMPISENGECLGRGIRVTLQDQPEIFLIQIRAMLRANADMGNLKILLPMWTSVEEID
EAKRLVDRAKEEVEQIAROI SMPQLGVMI E VPSMIFLLPOLKNRVDFVSGTNDLTQ
YLLAVDRNTHVASLYDCLHPAVIRALKLVFENSQAGLTISICGEMASFPFGALILI
GLGYSLSMSGCSVPRIKYLLRHLELAQVEKMAEILQAEISLRVKLSATFMEQHG
GGIVRGGI"

gene 10102. .10977

/gene="lgt"

/locus tag="plu0622"

10102. .10977

/gene="lgt"

/locus_tag="plu0622"

/codon_start=1

/transl_table=11

/product="Prolipoprotein diacylglyceryl transferase"

/protein_id="CAE12917.1"

/db_xref="GI:36784048"

/translation="MNSYLAFFENIDPVIFSGPIALHWYGFMYLVGFVFMWLATR
AAKPSGWTNKVENLLIYAGFVGVGGRIYVLFYFPAFLDNPLYLFLKVMDDGMSF
HGGLVGVICAMWFGRRTRKHFQVADFIAPLVPFGLGMGRIGNFINGELWGRVTLDT
PWAMLPSSRGEDIALAATDPSLLSLEQVLPVPRHPSQLYEMALEGIVLFIILNIYI
RKPRPMGVSGLFLIGVIGFRVIVEFRQPDQGLFDGISMGSQILSIPIMLAGILMW
IWAYKHQGNKVQEVK"
10964. .10969

gene

Query Match 2.0%; Score 24; DB 1; Length 342905;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGCCAAATAAAAAATGGCCATC 43

|||||

Db 27988 AGCCAAATAAAAAATGGCCATC 27965

RESULT 15

AX770908/c

LOCUS

AX770908 349980 bp DNA linear PAT 02-JUL-2003

DEFINITION Sequence 39 from Patent WO02094867.

ACCESSION AX770908

VERSION AX770908.1 GI:32438072

KEYWORDS

SOURCE

ORGANISM

Photorhabdus luminescens

Photorhabdus luminescens

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

Enterobacteriaceae; Photorhabdus.

1 Duchaud,E., Taourit,S., Glaser,P., Frangeul,L., Kunst,F.,

Danchin,A. and Buchrieser,C.

Sequence of the Photorhabdus luminescens strain TT01 genome and

uses

Patent: WO 02094867-A 39 28-NOV-2002;

INSTITUT PASTEUR (FR) ; CENTRE NATIONAL DE LA RECHERCHE

SCIENTIFIQUE (CNRS) (FR)

FEATURES

source

1. .349980

/location/Qualifiers

/organism="Photorhabdus luminescens"

/mol_type="genomic DNA"

/db_xref="taxon:29488"

/note="sequence length too big . Splitted into 3

sequences.-seq 00039: 1 to 349980-seq 10788: 300001 to

649980-seq 10789: 600001 to 700779"

ORIGIN

Query Match 2.0%; Score 24; DB 6; Length 349980;

Best Local Similarity 100.0%; Pred. No. 0.25;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGCCAAATAAAAAATGGCCATC 43

|||||

Db 122052 AGCCAAATAAAAAATGGCCATC 122029

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 12:32:16 ; Search time 728 Seconds
(without alignments)
9603.317 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctcgacagagatggaagaaaa.....ttttactttttttctgtgcag 1181

Scoring table:
OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

- Database : N Geneseqn16Dec04:*
- 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002as:*
 - 7: Geneseqn2002bs:*
 - 8: Geneseqn2003as:*
 - 9: Geneseqn2003bs:*
 - 10: Geneseqn2003cs:*
 - 11: Geneseqn2003ds:*
 - 12: Geneseqn2004as:*
 - 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1181	100.0	1181	3	Aaz36102 Nucleic a
2	31	2.6	31	3	Aaz36126 Primer de
3	31	2.6	31	3	Aaz36125 Primer de
4	24	2.0	2136	10	Acf69772 Phototab
5	24	2.0	110000	10	Continuation (28 o
6	24	2.0	110000	10	Continuation (2 of
7	22	1.9	22	3	Aaz36122 Primer de
8	22	1.9	22	3	Aaz36124 Primer de
9	22	1.9	22	3	Aaz36121 Primer de
10	21	1.8	21	3	Aaz36127 Primer de
11	21	1.8	341	8	Abx51595 Bovine ES
12	20	1.7	20	3	Aaz36123 Primer de
13	20	1.7	1077	6	Abq70353 Listeria
14	20	1.7	3589	2	Aax38295 Staphyloc
15	20	1.7	5402	4	Abll4236 Drosophil
16	20	1.7	19016	8	Aad47222 Streptoco
17	20	1.7	110000	6	ABQ69245 18
18	20	1.7	110000	6	ABA03041_17
19	20	1.7	319630	6	ABQ67194 Listeria
20	19	1.6	262	6	ABn76675 Human ORF

94	18	1.5	3526	13	ADS10272	Adsi0272 Human the	C 167	17	1.4	509	9	ADA06438	Human sec
95	18	1.5	3532	4	ABL29880	AbL29880 Drosophil	C 168	17	1.4	509	9	ADA39131	Human sec
96	18	1.5	3606	1	ADK65798	Adk65798 Angiogene	C 169	17	1.4	509	9	ADB96157	Human PRO
97	18	1.5	3686	11	ADM02994	Adm02994 Human cdn	C 170	17	1.4	509	10	ADC57629	Human PRO
98	18	1.5	4935	10	AAD51581	Aad51581 Human str	C 171	17	1.4	509	10	ADC54993	Human PRO
99	18	1.5	6126	10	ADK65846	Adk65846 Argiogene	C 172	17	1.4	509	10	ADC11860	Human sec
100	18	1.5	7315	13	ADR84254	Adr84254 Aspergill	C 173	17	1.4	509	10	ADC56282	Human PRO
101	18	1.5	8130	5	AAS34493	Aas34493 Human DNA	C 174	17	1.4	509	10	ADC07337	Human sec
102	18	1.5	8894	13	ADT05524	Adt05524 Haemophil	C 175	17	1.4	509	10	ADC11327	Human sec
103	18	1.5	19243	4	AAK71662	Aak71662 Human imm	C 176	17	1.4	509	10	ADC14449	Novel hum
104	18	1.5	26451	11	ACN44172	Acn44172 Mouse gen	C 177	17	1.4	509	10	ADD07981	Novel hum
105	18	1.5	40059	4	ABL15994	AbL15994 Drosophil	C 178	17	1.4	509	10	ADC81806	Human PRO
106	18	1.5	51469	4	AAK778813	Aak778813 Human imm	C 179	17	1.4	509	10	ADD07448	Novel hum
107	18	1.5	51469	4	AAK70270	Aak70270 Human imm	C 180	17	1.4	509	10	ADC82339	Human PRO
108	18	1.5	51469	4	AAK69322	Aak69322 Human imm	C 181	17	1.4	509	10	ADD08519	Novel hum
109	18	1.5	55216	9	AAI56825	Aai56825 Complete	C 182	17	1.4	509	10	ADD06768	Novel hum
110	18	1.5	55216	9	AAI56825	Aai56825 Complete	C 183	17	1.4	509	10	ADC83015	Human PRO
111	18	1.5	64184	13	ADS88580	Ads88580 Human hou	C 184	17	1.4	509	10	ADD55122	Human PRO
112	18	1.5	74037	6	ABK94412	Abk94412 DNA encod	C 185	17	1.4	509	10	ADD56080	Human PRO
113	18	1.5	79731	12	ADQ97640	Adq97640 Mouse can	C 186	17	1.4	509	10	ADD54518	Human PRO
114	18	1.5	83062	13	ABD32911_7	Abd32911_7 Continuation (8 of	C 187	17	1.4	509	10	AD266672	Novel hum
115	18	1.5	107304	13	ABD33230_0	Abd33230 Murine ca	C 188	17	1.4	509	10	AD261339	Novel hum
116	18	1.5	110000	2	AAT42063_04	Continuation (5 of	C 189	17	1.4	509	10	ADP67076	Human DNA
117	18	1.5	110000	10	ADG70184_0	Adg70184 DNA of BA	C 190	17	1.4	509	10	AD135330	Human PRO
118	18	1.5	110000	10	ACF42745_2	Continuation (3 of	C 191	17	1.4	509	10	ADH99822	Novel hum
119	18	1.5	117574	6	AAI45288	Aai45288 Human KCN	C 192	17	1.4	509	10	ABX77810	Human PRO
120	18	1.5	117576	13	ADT05736	Adt05736 Haemophil	C 193	17	1.4	509	10	ABX80222	Novel hum
121	18	1.5	128963	12	ADQ97110	Adq97110 Human can	C 194	17	1.4	509	10	ACA69128	Human EST
122	18	1.5	130877	13	ABD331108	Abd331108 Human can	C 195	17	1.4	509	10	ABX90199	Human sec
123	18	1.5	165554	11	ACN441108	Acn441108 Mouse gen	C 196	17	1.4	509	10	ABX64045	Human PRO
124	18	1.5	198073	11	ACN44302	Acn44302 Human gen	C 197	17	1.4	509	12	ADP35275	Human PRO
125	18	1.5	236303	4	AAS11614	Aas11614 Human gen	C 198	17	1.4	509	12	ADG11525	Human PRO
126	18	1.5	260803	13	ABD32730	Abd32730 Human can	C 199	17	1.4	509	12	ADH19395	Human sec
127	17	1.4	172	10	ADB67981	Adb67981 Human lun	C 200	17	1.4	509	12	ADH20888	Human sec
128	17	1.4	219	2	AAT21771	Aat21771 Human gen	C 201	17	1.4	509	12	ADH19928	Human sec
129	17	1.4	343	4	AAS41614	Aas41614 cDNA encd	C 202	17	1.4	520	13	ACN58485	AcN58485 Cotton gy
130	17	1.4	349	6	ABQ85850	Abq85850 Arabidops	C 203	17	1.4	520	13	ACN58485	AcN58485 Cotton gy
131	17	1.4	350	10	ADA19244	Ada19244 Human ins	C 204	17	1.4	547	13	ADQ55319	Novel can
132	17	1.4	353	4	ABK42911	Abk42911 Genomic s	C 205	17	1.4	550	9	ACH42322	Human foe
133	17	1.4	353	4	ABK42910	Abk42910 Genomic s	C 206	17	1.4	550	9	ACH42322	Human foe
134	17	1.4	353	9	ADB61066	AdB61066 Connectiv	C 207	17	1.4	606	11	ABD16686	Pseudomon
135	17	1.4	353	9	ADB61067	AdB61067 Connectiv	C 208	17	1.4	651	6	ABN67448	Abn67448 Streptoco
136	17	1.4	380	3	ADF57045	Adf57045 Urogenita	C 209	17	1.4	652	6	ABQ60654	Abq60654 Human col
137	17	1.4	391	5	ABV07635	Abv07635 Human pro	C 210	17	1.4	659	4	AAI91805	Aai91805 Human pol
138	17	1.4	457	4	AAI14328	Aai14328 Human bre	C 211	17	1.4	678	11	ABD15038	Abd15038 Pseudomon
139	17	1.4	489	6	ABN66862	Abn66862 Streptoco	C 212	17	1.4	733	13	ADR13134	Adr13134 Human can
140	17	1.4	493	5	AAF93217	Aaf93217 cDNA encd	C 213	17	1.4	859	6	ABN98833	Abn98833 Arabidops
141	17	1.4	507	12	ADQ21034	Adq21034 Human sof	C 214	17	1.4	913	4	AAD08247	Aad08247 Rat inter
142	17	1.4	509	3	AAZ64994	Aaz64994 Membrane-	C 215	17	1.4	942	11	ABD16842	Abd16842 Pseudomon
143	17	1.4	509	3	AAA44926	Aaa44926 Human sec	C 216	17	1.4	969	9	ADB11831	Adb11831 Alloiococ
144	17	1.4	509	5	AAF44140	Aaf44140 Human EST	C 217	17	1.4	969	9	ADB11833	Adb11833 Alloiococ
145	17	1.4	509	8	ACA64267	AcA64267 Human sec	C 218	17	1.4	981	6	ABK35066	Abk35066 Human CDN
146	17	1.4	509	8	ABX80726	Abx80726 Human sec	C 219	17	1.4	987	8	ACA23210	AcA23210 Prokaryot
147	17	1.4	509	8	ACD44235	AcD44235 Human PRO	C 220	17	1.4	1001	3	AAH51553	Aah51553 Human CYP
148	17	1.4	509	8	ABX79406	Abx79406 Human sec	C 221	17	1.4	1044	3	ACA54061	AcA54061 Arabidops
149	17	1.4	509	8	ACA93427	AcA93427 Novel hum	C 222	17	1.4	1077	8	ACF39368	AcF39368 Mycobacte
150	17	1.4	509	8	ABX81109	Abx81109 Novel hum	C 223	17	1.4	1082	4	AAS41078	Aas41078 cDNA enco
151	17	1.4	509	8	ACA92925	AcA92925 Novel hum	C 224	17	1.4	1089	11	ABD02893	Abd02893 Pseudomon
152	17	1.4	509	8	ABX17009	Abx17009 Human PRO	C 225	17	1.4	1094	6	ABQ69053	Abq69053 Listeria
153	17	1.4	509	9	ACA67864	AcA67864 Novel hum	C 226	17	1.4	1110	3	AAC41607	Aac41607 Arabidops
154	17	1.4	509	9	ACA88313	AcA88313 Human sec	C 227	17	1.4	1110	11	ABD16948	Abd16948 Pseudomon
155	17	1.4	509	9	ACD81820	AcD81820 Human PRO	C 228	17	1.4	1113	3	AAC33438	Aac33438 Arabidops
156	17	1.4	509	9	ADA37660	Ada37660 Human sec	C 229	17	1.4	1193	6	AAD31839	Aad31839 Human pan
157	17	1.4	509	9	ADA21346	Ada21346 Human sec	C 230	17	1.4	1203	8	ACF74423	AcF74423 Staphyloc
158	17	1.4	509	9	ADA10133	Ada10133 Human sec	C 231	17	1.4	1245	13	ADS61566	AdS61566 Bacteriol
159	17	1.4	509	9	ADA17677	Ada17677 Human PRO	C 232	17	1.4	1325	8	ACA48583	AcA48583 Prokaryot
160	17	1.4	509	9	ADA27785	Ada27785 Human sec	C 233	17	1.4	1325	6	ABK48731	Abk48731 DNA encod
161	17	1.4	509	9	ADA94365	Ada94365 Human sec	C 234	17	1.4	1328	10	ACF04930	AcF04930 plant wit
162	17	1.4	509	9	ADA38590	Ada38590 Human sec	C 235	17	1.4	1328	12	ADP90953	Adp90953 Figleaf g
163	17	1.4	509	9	ADA92711	Ada92711 Human sec	C 236	17	1.4	1328	13	ADR38363	Adr38363 Fig leaf
164	17	1.4	509	9	ACH65381	Ach65381 Human sec	C 237	17	1.4	1338	12	ADN74210	Adn74210 Thale cre
165	17	1.4	509	9	ADA22272	Ada22272 Human sec	C 238	17	1.4	1359	11	ABD15193	Abd15193 Pseudomon
166	17	1.4	509	9	ACD39371	AcD39371 Human cdn	C 239	17	1.4	1389	11	ABD16642	Abd16642 Pseudomon

c 240	17	1.4	1443	13	ADS58627	AdS58627 Bacterial	313	17	1.4	6643	12	ADQ83730	AdQ83730 Human tum
241	17	1.4	1446	5	AAS15898	Aas15898 DNA encod	314	17	1.4	7075	2	AAV74574	AAV74574 Staphyloc
242	17	1.4	1656	11	ABD14941	Abd14941 Peedomon	c 315	17	1.4	7737	4	ABK43114	ABk43114 Genomic s
c 243	17	1.4	1692	13	ADT44408	Adt44408 Bacterial	c 316	17	1.4	7737	4	AAK90942	AAk90942 Human dig
244	17	1.4	1704	4	AAI60181	Aai60181 Human pol	c 317	17	1.4	7737	5	AAS31977	Aas31977 Human liv
c 245	17	1.4	1708	4	AAI58395	Aai58395 Human pol	c 318	17	1.4	7737	5	ABN90332	ABn90332 Human liv
c 246	17	1.4	1708	5	ADQ98605	Adq98605 DNA encod	c 319	17	1.4	7737	9	ADB61270	ADB61270 Connectiv
c 247	17	1.4	1708	9	ADB48365	Adb48365 Novel hum	c 320	17	1.4	7737	11	ADJ15245	Adj15245 Human liv
c 248	17	1.4	1710	4	AAK51944	Aak51944 Human pol	c 321	17	1.4	9614	4	ABL03154	ABl03154 Drosophil
c 249	17	1.4	1715	4	ABA09030	AbA09030 Human sec	c 322	17	1.4	9771	4	AAK71552	AAk71552 Human imm
c 250	17	1.4	1715	4	AAK52928	Aak52928 Human pol	c 323	17	1.4	10669	4	AAK81696	AAk81696 Human imm
c 251	17	1.4	1722	4	AAF83199	Aaf83199 Nucleotid	c 324	17	1.4	10669	4	AAK76914	AAk76914 Human imm
c 252	17	1.4	1722	12	ADL97962	Adl97962 Human nel	c 325	17	1.4	11314	4	ABL06868	ABl06868 Drosophil
c 253	17	1.4	1736	4	AALH5307	AalH5307 Human cdn	c 326	17	1.4	11446	4	AAK90464	AAk90464 Human dig
c 254	17	1.4	1770	11	ABD02834	Abd02834 Peedomon	c 327	17	1.4	12258	12	ADO42288	ADO42288 Human NOV
c 255	17	1.4	1815	8	ABV72425	Abv72425 Nucleotid	c 328	17	1.4	13971	10	ADE78571	Ade78571 Dynein ax
c 256	17	1.4	1902	6	ABQ70018	Abq70018 Listeria	c 329	17	1.4	14111	4	ABL10392	ABl10392 Drosophil
c 257	17	1.4	1902	6	ABQ68241	Abq68241 Listeria	c 330	17	1.4	14111	4	AAS57100	Aas57100 DNA encod
c 258	17	1.4	1950	10	ABT42398	Abt42398 Toxicity	c 331	17	1.4	14111	10	ADC35791	AdC35791 Drosophil
c 259	17	1.4	2013	11	ADQ81940	Adq81940 Human sig	c 332	17	1.4	14676	4	AAK71553	AAk71553 Human imm
c 260	17	1.4	2068	11	ADM02774	Adm02774 Human cdn	c 333	17	1.4	15115	4	ABL13702	ABl13702 Drosophil
c 261	17	1.4	2181	8	ACA51023	Aca51023 Prokaryot	c 334	17	1.4	15141	4	ABL19378	ABl19378 Drosophil
c 262	17	1.4	2181	8	ACA51893	Aca51893 Prokaryot	c 335	17	1.4	15479	6	ABK39965	ABk39965 Human che
c 263	17	1.4	2187	4	ABL17227	AbL17227 Drosophil	c 336	17	1.4	15496	4	AAK71549	AAk71549 Human imm
c 264	17	1.4	2274	12	ADP98783	Adp98783 C. albica	c 337	17	1.4	18705	12	ADJ12609	Adj12609 DNA fragm
c 265	17	1.4	2297	4	AAH33960	Aah33960 Human cdn	c 338	17	1.4	18715	12	ADJ12533	AdJ12533 DNA fragm
c 266	17	1.4	2297	12	ADQ84343	Adq84343 Human tum	c 339	17	1.4	18830	4	AAK71558	AAk71558 Human imm
c 267	17	1.4	2297	13	ADQ87243	Adq87243 Human tum	c 340	17	1.4	23823	4	AAK79160	AAk79160 Human imm
c 268	17	1.4	2297	13	ACN38292	Acn38292 Tumour-as	c 341	17	1.4	23823	4	AAK79161	AAk79161 Human imm
c 269	17	1.4	2332	5	ADM19295	Adm19295 Novel hum	c 342	17	1.4	25911	4	ABL02604	ABl02604 Drosophil
c 270	17	1.4	2346	3	AAK77834	Aak77834 Human can	c 343	17	1.4	25929	4	ABL02602	ABl02602 Drosophil
c 271	17	1.4	2357	5	ADM19553	Adm19553 Novel hum	c 344	17	1.4	26502	12	ADL18577	AdL18577 Human wil
c 272	17	1.4	2362	6	AAD24417	Aad24417 Human RNA	c 345	17	1.4	26556	10	ACN44944	AcN44944 Mouse gen
c 273	17	1.4	2391	8	ACA23469	Aca23469 Prokaryot	c 346	17	1.4	26750	11	ACN44944	AcN44944 Mouse gen
c 274	17	1.4	2403	13	ADT44287	Adt44287 Bacterial	c 347	17	1.4	28871	13	ADT05539	Adt05539 Haemophil
c 275	17	1.4	2430	2	AAV01880	Aav01880 Human nel	c 348	17	1.4	29596	4	ABL06866	ABl06866 Drosophil
c 276	17	1.4	2430	8	ACC69971	Acc69971 Human nel	c 349	17	1.4	32216	4	AAK06925	AAk06925 Human rep
c 277	17	1.4	2430	13	ADQ76514	Adq76514 Nucleotid	c 350	17	1.4	32216	4	ABA07741	ABa07741 Human ova
c 278	17	1.4	2433	13	ADR32022	Adr32022 Human NEL	c 351	17	1.4	32217	4	AAK41742	AAa41742 Genomic s
c 279	17	1.4	2453	5	AAS90752	Aas90752 DNA encod	c 352	17	1.4	33012	8	ABSS5899	ABss5899 Bovine ad
c 280	17	1.4	2478	11	ACH99175	Ach99175 Klebsiell	c 353	17	1.4	33113	8	ABSS5900	ABss5900 Bovine ad
c 281	17	1.4	2646	6	ABK93186	ABk93186 Human pro	c 354	17	1.4	33306	8	ABSS5901	ABss5901 Bovine ad
c 282	17	1.4	2894	10	ADRS9174	Adrs9174 Human gen	c 355	17	1.4	33310	8	ABSS5903	ABss5903 Bovine ad
c 283	17	1.4	2906	10	ADG37196	Adg37196 Nuclear f	c 356	17	1.4	34063	9	ACD19083	ACd19083 E. coli 0
c 284	17	1.4	2910	6	ABQ84763	Abq84763 Chlamydia	c 357	17	1.4	34079	3	ABSS5902	ABss5902 Bovine ad
c 285	17	1.4	2910	6	ABQ78021	Abq78021 Chlamydia	c 358	17	1.4	34185	3	AAC62130	AAc62130 Nucleotid
c 286	17	1.4	2910	8	ABX99188	Abx99188 C. psitta	c 359	17	1.4	34185	8	ABSS5888	ABss5888 Bovine ad
c 287	17	1.4	2977	2	AAV01881	Aav01881 Human nel	c 360	17	1.4	34185	10	ADC17118	Adc17118 Bovine ad
c 288	17	1.4	2977	4	AAF93198	Aaf93198 Human NEL	c 361	17	1.4	39071	12	ADM97420	Adm97420 Prostate
c 289	17	1.4	2977	6	ABL65425	AbL65425 Lung canc	c 362	17	1.4	39703	10	ADC86430	AdC86430 Human GPC
c 290	17	1.4	2977	8	ACC69972	Acc69972 Human nel	c 363	17	1.4	44029	10	ADC00710	AdC00710 Enterohae
c 291	17	1.4	2977	10	ACC72855	Acc72855 Human can	c 364	17	1.4	45175	9	ACD19103	ACd19103 E. coli 0
c 292	17	1.4	2977	12	ADL97961	Adl97961 Human NEL	c 365	17	1.4	48509	13	ABD33397	ABd33397 Human can
c 293	17	1.4	2977	12	ADQ21159	Adq21159 Human sof	c 366	17	1.4	49175	13	ABD32540	ABd32540 Human can
c 294	17	1.4	2977	13	ADQ76515	Adq76515 Nucleotid	c 367	17	1.4	49650	10	ADC00365	AdC00365 Enterohae
c 295	17	1.4	3024	2	AAT44120	Aat44120 Rat brain	c 368	17	1.4	49999	2	AAZ23900	AAz23900 Human LOB
c 296	17	1.4	3096	12	ADL28610	Adl28610 Mycoplasma	c 369	17	1.4	49999	2	AAZ23896	AAz23896 Murine LO
c 297	17	1.4	3156	4	AAK84100	Aak84100 Human imm	c 370	17	1.4	49999	2	AAZ23891	AAz23891 Murine LO
c 298	17	1.4	3240	11	ACH94536	Ach94536 Klebsiell	c 371	17	1.4	54391	10	ACF42745	ACf42745 Contuinuati (4 of
c 299	17	1.4	3277	13	ADR07313	Adr07313 Pull leng	c 372	17	1.4	59560	4	AAK80620	AAk80620 Human imm
c 300	17	1.4	3297	4	AAF94453	Aaf94453 Human hyd	c 373	17	1.4	59560	4	AAK71550	AAk71550 Human imm
c 301	17	1.4	3329	13	ADR07601	Adr07601 Full leng	c 374	17	1.4	60940	9	ADA02582	ADa02582 Human REL
c 302	17	1.4	3486	12	ADQ96439	Adq96439 T cell ac	c 375	17	1.4	60940	10	ADB72320	ABd72320 Human REL
c 303	17	1.4	3653	6	ABQ70935	Abq70935 Listeria	c 376	17	1.4	60940	10	ADE95830	ADe95830 Human REL
c 304	17	1.4	3747	4	ABL02690	ABl02690 Drosophil	c 377	17	1.4	73102	12	ADQ97888	ADq97888 Human can
c 305	17	1.4	3940	10	ADRS3669	Adrs3669 Human pro	c 378	17	1.4	81001	4	AAF30035	Aaf30035 Human apo
c 306	17	1.4	4052	12	ADM931284	Adm931284 DNA homol	c 379	17	1.4	83009	13	ABD32789	ABd32789 Mouse can
c 307	17	1.4	4645	10	ADBS3601	Adbs3601 Primary r	c 380	17	1.4	83400	12	ADP07906	ADp07906 Human RAD
c 308	17	1.4	5356	4	AAK68208	Aak68208 Human imm	c 381	17	1.4	91740	10	ADC00956	AdC00956 Enterohae
c 309	17	1.4	5541	4	ABL17226	ABl17226 Drosophil	c 382	17	1.4	94001	13	ADJ33491	ADj33491 Human LAR
c 310	17	1.4	6101	13	ADS89077	AdS89077 Human CYP	c 383	17	1.4	96960	8	ACF62734	ACf62734 Cancer ba
c 311	17	1.4	6118	6	ABL33031	ABl33031 Human imm	c 384	17	1.4	96960	8	ADB20849	ABd20849 MRP1 base
c 312	17	1.4	6551	4	ABL13538	ABl13538 Drosophil	c 385	17	1.4	96960	10	ADB87938	ABd87938 Human UGT

386	17	1.4	96960	10	ADB96921	Adb96921 Human MDR	459	16	1.4	121	12	ADK92429	Adk92429 Polynucle
387	17	1.4	96960	10	ADB92112	Adb92112 Human MDR	460	16	1.4	131	8	AAU24874	Aal24874 Human bre
388	17	1.4	110000	2	AAT42063_05	Continuation (6 of	C 461	16	1.4	162	8	ACA36987	Aac36987 Prokaryot
389	17	1.4	110000	2	AAV30458_4	Continuation (5 of	C 462	16	1.4	166	2	AAV88439	Aav88439 EST clone
390	17	1.4	110000	2	AAV30459_4	Continuation (5 of	C 463	16	1.4	181	4	AAU21926	Aal21926 Human bre
391	17	1.4	110000	2	AAX20248_00	Aax20248 Borrelia	464	16	1.4	182	4	AAU24773	Aal24773 Human bre
392	17	1.4	110000	2	AAZ01425_08	Continuation (9 of	465	16	1.4	186	10	ADK11593	Adk11593 Breast ca
393	17	1.4	110000	3	AAF22305_08	Continuation (9 of	C 466	16	1.4	201	13	ADS36689	Ads36689 Human aut
394	17	1.4	110000	4	AAI99682_07	Continuation (8 of	467	16	1.4	201	13	ADS40347	Ads40347 Human aut
395	17	1.4	110000	4	AAI99682_21	Continuation (22 o	468	16	1.4	224	9	ADA01570	Ada01570 Mouse Gfi
396	17	1.4	110000	4	AAI99682_22	Continuation (23 o	469	16	1.4	224	10	ADB71309	Adb71309 Mouse car
397	17	1.4	110000	4	AAI99682_26	Continuation (27 o	470	16	1.4	228	2	AAH86704	Aah86704 Human sin
398	17	1.4	110000	4	AAI99682_38	Continuation (39 o	471	16	1.4	233	12	ACH90714	Ach90714 Human gen
399	17	1.4	110000	4	AAI99683_07	Continuation (8 of	472	16	1.4	237	4	AAI51840	Aai51840 Probe #20
400	17	1.4	110000	4	AAI99683_21	Continuation (22 o	473	16	1.4	237	4	AAK45924	Aak45924 Human bon
401	17	1.4	110000	4	AAI99683_26	Continuation (27 o	474	16	1.4	237	4	ABS45622	Abs45622 Human liv
402	17	1.4	110000	6	ABN71527_18	Continuation (19 o	475	16	1.4	237	6	ABS20211	Abs20211 Human gen
403	17	1.4	110000	6	ABN71527_19	Continuation (20 o	C 476	16	1.4	247	6	ABS72896	Abs72896 Human gen
404	17	1.4	110000	6	ABA03041_22	Continuation (23 o	C 477	16	1.4	249	4	AAS53128	Aas53128 Enterococ
405	17	1.4	110000	9	ADB12064_13	Continuation (14 o	478	16	1.4	264	10	ADC91677	Adc91677 E. faeciu
406	17	1.4	110000	10	ADP77343_17	Continuation (18 o	C 479	16	1.4	266	4	AAL22093	Aal22093 Human bre
407	17	1.4	110000	10	ADG70447_0	Adg70447 Human ANG	C 480	16	1.4	270	4	AAS51383	Aas51383 Enterococ
408	17	1.4	110000	10	ABZ79565_0	Abz79565 CLLD8 and	481	16	1.4	273	11	ACH97138	Ach97138 Klebsiell
409	17	1.4	110000	12	AD034435_1	Continuation (2 of	C 482	16	1.4	274	9	ADA01571	Ada01571 Mouse Gfi
410	17	1.4	110000	12	AD079173_3	Continuation (4 of	C 483	16	1.4	274	10	ADB71310	Adb71310 Mouse car
411	17	1.4	110000	12	AD097331_0	Adg97331 Human can	484	16	1.4	275	4	AAL13178	Aal13178 Human bre
412	17	1.4	110000	13	ABD32627_1	Continuation (2 of	C 485	16	1.4	275	4	AAL13096	Aal13096 Human bre
413	17	1.4	112132	6	ABK90888	Abk90888 Human ATP	486	16	1.4	275	6	ABL38371	AbL38371 Human col
414	17	1.4	112132	11	ADM56267	Adm56267 Human ATP	C 487	16	1.4	278	7	ADS71806	Ads71806 Human kid
415	17	1.4	123785	10	ABX77171	Abx77171 DNA seque	488	16	1.4	279	9	ADA01519	Ada01519 Mouse Gfi
416	17	1.4	132197	5	ACA61370	Aai61370 Soybean 5	C 489	16	1.4	294	10	ADB71258	Adb71258 Mouse car
417	17	1.4	132492	8	ACA64895	Aca64895 Human GAB	C 490	16	1.4	294	4	AAK60547	Aak60547 Human imm
418	17	1.4	135800	11	ACNA4446	Acn44446 Human gen	C 491	16	1.4	295	4	AAL13223	Aal13223 Human bre
419	17	1.4	139389	6	ABK84795	Abk84795 Human CDN	C 492	16	1.4	297	4	AAL14527	Aal14527 Human bre
420	17	1.4	139389	9	ADB70369	Adb70369 PAC 6802	C 493	16	1.4	298	10	ABT22420	Abt22420 Breast ca
421	17	1.4	139389	12	ADJ37148	Adj37148 Human mal	494	16	1.4	300	2	AAZ14182	Aaz14182 Human gen
422	17	1.4	144792	10	ADC87620	Adc87620 Human GPC	495	16	1.4	303	4	AAL24772	Aal24772 Human bre
423	17	1.4	154394	12	ADQ17784	Adq17784 Human sof	496	16	1.4	303	4	AAL21928	Aal21928 Human bre
424	17	1.4	154465	6	AAD28763	Aad28763 Human AKA	C 497	16	1.4	303	6	ABL64724	AbL64724 Lung canc
425	17	1.4	158245	6	AAD28762	Aad28762 Human AKA	C 498	16	1.4	303	6	ABL65680	AbL65680 Lung canc
426	17	1.4	161425	4	AAH02340	Aah02340 Human AKA	C 499	16	1.4	305	4	AAL24900	Aal24900 Human bre
427	17	1.4	162025	4	AAH02339	Aah02339 Human AKA	C 500	16	1.4	305	4	AAL24885	Aal24885 Human bre
428	17	1.4	162025	6	AAD28758	Aad28758 Human AKA	501	16	1.4	305	4	AAL24909	Aal24909 Human bre
429	17	1.4	162025	6	AAD28759	Aad28759 Human AKA	502	16	1.4	305	4	AAL24865	Aal24865 Human bre
430	17	1.4	162025	13	ADS75958	Ads75958 Human A-k	503	16	1.4	306	4	AAL16030	Aal16030 Human bre
431	17	1.4	162025	13	ADS75959	Ads75959 Human A-k	C 504	16	1.4	306	4	AAL04236	Aal04236 Human rep
432	17	1.4	162025	13	ADS75963	Ads75963 Human A-k	C 505	16	1.4	307	12	ADM42828	Adm42828 DNA encod
433	17	1.4	162025	13	ADS75962	Ads75962 Human A-k	C 506	16	1.4	308	3	ACS93814	Acs93814 Cat flea
434	17	1.4	168325	11	ACNA4484	Acn44484 Mouse gen	507	16	1.4	308	4	AAL13056	Aal13056 Human bre
435	17	1.4	175590	10	ADD50650	Add50650 BAC seque	508	16	1.4	308	4	AAL04235	Aal04235 Human rep
436	17	1.4	188971	12	ADL08108	Adl08108 Human gen	C 509	16	1.4	309	4	AAL04235	Aal04235 Human rep
437	17	1.4	260027	11	ACNA4046	Acn44046 Human gen	C 510	16	1.4	310	4	AAL13142	Aal13142 Human bre
438	17	1.4	277616	13	ABD32602	Abd32602 Human can	C 511	16	1.4	310	12	ADL09936	AdL09936 Cat flea
439	17	1.4	289190	13	ABD33143	Abd33143 Murine can	512	16	1.4	313	4	AAL13058	Aal13058 Human bre
440	17	1.4	295036	11	ACNA4068	Acn44068 Mouse gen	C 513	16	1.4	313	4	AAL00799	Aal00799 Human rep
441	17	1.4	304326	13	ADSL15253	Adsl15253 Rat senso	C 514	16	1.4	314	4	AAL13111	Aal13111 Human bre
442	17	1.4	308766	13	ADT05738	Adt05738 Haemophil	C 515	16	1.4	314	4	AAL14409	Aal14409 Human bre
443	17	1.4	326014	6	ABK89296	Abk89296 Human gen	C 516	16	1.4	315	4	AAL22027	Aal22027 Human bre
444	17	1.4	326014	12	ADQ94981	Adq94981 Human kin	C 517	16	1.4	316	4	AAL22054	Aal22054 Human bre
445	16	1.4	18	10	ADEL13446	Adel13446 HLA class	C 518	16	1.4	316	4	AAL22084	Aal22084 Human bre
446	16	1.4	18	10	ADEL7633	Adel7633 Human pro	C 519	16	1.4	316	4	AAL13157	Aal13157 Human bre
447	16	1.4	18	12	ADL09296	Adl09296 HLA locus	C 520	16	1.4	316	4	AAL13214	Aal13214 Human bre
448	16	1.4	25	9	ACI71716	AcI71716 Human mic	521	16	1.4	316	10	ABT22417	Abt22417 Breast ca
449	16	1.4	29	10	ACF79627	Acf79627 Thiopurin	C 522	16	1.4	317	4	AAL22308	Aal22308 Human bre
450	16	1.4	33	6	AAI71477	Aai71477 Human pax	C 523	16	1.4	317	4	AAL22012	Aal22012 Human bre
451	16	1.4	35	4	AAF83799	Aaf83799 Dengue-3	C 524	16	1.4	317	9	ADA00298	Ada00298 Human alp
452	16	1.4	39	3	AAZ54730	Aaz54730 Neisseria	C 525	16	1.4	317	10	ABT22452	Abt22452 Breast ca
453	16	1.4	39	3	AAZ54726	Aaz54726 Neisseria	526	16	1.4	317	10	ABT22411	Abt22411 Breast ca
454	16	1.4	60	2	AAQ79629	Aaq79629 Probe EML	527	16	1.4	320	4	AAL22048	Aal22048 Human bre
455	16	1.4	65	6	ABN55374	Abn55374 Mouse spl	528	16	1.4	320	4	AAL13177	Aal13177 Human bre
456	16	1.4	115	3	AAC24930	Aac24930 Human sec	C 529	16	1.4	320	4	AAL13184	Aal13184 Human bre
457	16	1.4	115	12	ADG74045	Adg74045 Amyloid b	530	16	1.4	320	4	AAL22047	Aal22047 Human bre
458	16	1.4	117	12	ADG74043	Adg74043 Amyloid b	531	16	1.4	320	10	ABT22451	Abt22451 Breast ca

C 532	16	1.4	321	4	AAL13200	Aal13200 Human bre	605	16	1.4	428	13	ACN49875	Acn49875 Cotton pr
C 533	16	1.4	321	4	AAL21966	Aal21966 Human bre	C 607	16	1.4	429	9	ACH17696	Ach17696 Human adu
C 534	16	1.4	321	4	AAL22070	Aal22070 Human bre	C 606	16	1.4	434	4	AAL14505	Aal14505 Human bre
C 535	16	1.4	321	11	ACH97923	Ach97923 Klebsiell	C 608	16	1.4	437	3	AAZ34746	Aaz34746 Mouse ant
C 536	16	1.4	324	6	ABK79959	Abk79959 Bacillus	C 609	16	1.4	439	4	AAK62904	Aak62904 Human imm
C 537	16	1.4	327	4	AAL14510	Aal14510 Human bre	C 610	16	1.4	444	4	AAL14410	Aal14410 Human bre
C 538	16	1.4	327	4	AAL24894	Aal24894 Human bre	C 611	16	1.4	448	4	AAL15009	Aal15009 Probe #49
C 539	16	1.4	328	4	AAL23372	Aal23372 Human bre	C 612	16	1.4	448	4	ABA56745	Aba56745 Human foe
C 540	16	1.4	332	4	AAL21956	Aal21956 Human bre	C 613	16	1.4	448	4	AAI36353	AAi36353 Probe #50
C 541	16	1.4	335	4	AAL21970	Aal21970 Human bre	C 614	16	1.4	448	4	ABA46208	Aba46208 Human bre
C 542	16	1.4	335	6	ABL79008	Ab179008 Human ova	C 615	16	1.4	448	4	ABA26372	Aba26372 Probe #48
C 543	16	1.4	336	4	AAI26038	AAi26038 Probe #15	C 616	16	1.4	448	4	AAK30400	Aak30400 Human bon
C 544	16	1.4	336	4	ABA73200	Aba73200 Human foe	C 617	16	1.4	448	4	AAK04875	Aak04875 Human bra
C 545	16	1.4	336	4	AAI53628	AAi53628 Probe #22	C 618	16	1.4	448	4	ABS30040	Abs30040 Human liv
C 546	16	1.4	336	4	ABA38635	Aba38635 Probe #17	C 619	16	1.4	448	5	AAI04776	AAi04776 Probe #47
C 547	16	1.4	336	4	AAK47800	Aak47800 Human bon	C 620	16	1.4	448	6	ABS05007	Abs05007 Human gen
C 548	16	1.4	336	4	AAK21635	Aak21635 Human bra	C 621	16	1.4	450	11	ACH95286	Ach95286 Klebsiell
C 549	16	1.4	336	4	ABA547525	Aba547525 Human liv	C 622	16	1.4	450	12	ADQ03102	Adq03102 Klebsiell
C 550	16	1.4	336	6	ABS21808	Abs21808 Human gen	C 623	16	1.4	455	9	ACH28680	Ach28680 Human adu
C 551	16	1.4	339	10	ACD97273	Acd97273 Human col	C 624	16	1.4	456	4	AAK88148	Aak88148 Human dig
C 552	16	1.4	339	10	ABT22449	Abt22449 Breat ca	C 625	16	1.4	457	10	ADK15270	Adk15270 Urinary s
C 553	16	1.4	340	2	AAV09791	AAv09791 DNA encod	C 626	16	1.4	457	10	ADK15271	Adk15271 Urinary s
C 554	16	1.4	340	4	AAL23272	Aal23272 Human bre	C 627	16	1.4	459	4	AAI14446	AAi14446 Human foe
C 555	16	1.4	346	11	ACN91050	Acn91050 Breat ca	C 628	16	1.4	461	4	ABA57016	Aba57016 Human foe
C 556	16	1.4	348	4	AAL13086	Aal13086 Human bre	C 629	16	1.4	461	10	ACD95241	Acd95241 Human col
C 557	16	1.4	349	10	ABT22421	Abt22421 Breat ca	C 630	16	1.4	463	5	ADL43062	Adl43062 Human ova
C 558	16	1.4	350	6	ABK45438	Abk45438 cDNA enco	C 631	16	1.4	466	4	AAI23207	AAi23207 Human bre
C 559	16	1.4	351	10	ABT22450	Abt22450 Breat ca	C 632	16	1.4	468	9	ACH45436	Ach45436 Human foe
C 560	16	1.4	353	12	ADL87583	Adl87583 DNA up-re	C 633	16	1.4	469	4	AAH12542	AAh12542 Human CDN
C 561	16	1.4	353	12	ADL87584	Adl87584 DNA up-re	C 634	16	1.4	470	10	ABT22511	Abt22511 Breat ca
C 562	16	1.4	354	4	AAL16056	Aal16056 Human bre	C 635	16	1.4	474	4	AAI14345	AAi14345 Human bre
C 563	16	1.4	355	4	AAL16050	Aal16050 Human bre	C 636	16	1.4	475	12	ADP95480	Adp95480 Cotton ex
C 564	16	1.4	355	4	AAL15929	Aal15929 Human bre	C 637	16	1.4	476	4	AAI23388	AAi23388 Human bre
C 565	16	1.4	356	4	AAL15928	Aal15928 Human bre	C 638	16	1.4	476	4	AAI23303	AAi23303 Human bre
C 566	16	1.4	358	4	AAI16041	AAi16041 Human bre	C 639	16	1.4	477	4	AAI23367	AAi23367 Human bre
C 567	16	1.4	359	4	AAI38954	AAi38954 Novel hum	C 640	16	1.4	478	11	ACN90710	Acn90710 Breat ca
C 568	16	1.4	360	4	AAL16065	Aal16065 Human bre	C 641	16	1.4	479	4	AAI14441	AAi14441 Human bre
C 569	16	1.4	368	2	AAV09793	AAv09793 DNA encod	C 642	16	1.4	480	10	ADE82183	Ade82183 Arabidops
C 570	16	1.4	368	2	AAV09789	AAv09789 DNA encod	C 643	16	1.4	481	4	AAI23341	AAi23341 Human bre
C 571	16	1.4	376	4	AAI81022	AAi81022 Human pol	C 644	16	1.4	481	13	ADQ57208	Adq57208 Novel can
C 572	16	1.4	376	8	ACC55387	Acc55387 Rice endo	C 645	16	1.4	485	6	ABL79973	Ab179973 Human ova
C 573	16	1.4	377	2	AAV89818	AAv89818 EST clone	C 646	16	1.4	485	9	ACH41429	Ach41429 Human foe
C 574	16	1.4	392	10	ADE09780	Ade09780 Novel DNA	C 647	16	1.4	489	4	AAI12845	AAi12845 Probe #27
C 575	16	1.4	396	9	ADA29411	Ada29411 DNA encod	C 648	16	1.4	489	4	ABA54548	Aba54548 Human foe
C 576	16	1.4	398	5	AAF64550	Aaf64550 Novel hum	C 649	16	1.4	489	4	AAI34201	AAi34201 Probe #28
C 577	16	1.4	399	6	ABL63988	Ab163988 Breat ca	C 650	16	1.4	489	4	ABA44091	Aba44091 Human bre
C 578	16	1.4	399	6	ABL63569	Ab163569 Breat ca	C 651	16	1.4	489	4	ABA24330	Aba24330 Probe #27
C 579	16	1.4	400	6	ABL82366	Ab182366 Human ova	C 652	16	1.4	489	4	AAK28280	Aak28280 Human bon
C 580	16	1.4	400	11	ACN88419	Acn88419 Breat ca	C 653	16	1.4	489	4	AAK02838	Aak02838 Human bra
C 581	16	1.4	401	5	AAF64548	Aaf64548 Novel hum	C 654	16	1.4	489	4	ABS27884	Abs27884 Human liv
C 582	16	1.4	404	4	AAI82297	AAi82297 Human pol	C 655	16	1.4	489	5	AAI02762	AAi02762 Probe #27
C 583	16	1.4	405	4	AAI86263	AAi86263 Human pol	C 656	16	1.4	489	6	ABS02793	Abs02793 Human gen
C 584	16	1.4	405	8	ABX54458	Abx54458 Bovine ES	C 657	16	1.4	491	4	AAI23316	AAi23316 Human bre
C 585	16	1.4	405	9	ACH30624	Ach30624 Human tes	C 658	16	1.4	492	4	AAI24880	AAi24880 Human bre
C 586	16	1.4	409	9	ACH18686	Ach18686 Human adu	C 659	16	1.4	494	6	ABQ60169	Abq60169 Human col
C 587	16	1.4	414	10	ABT22508	Abt22508 Breat ca	C 660	16	1.4	494	9	ACH35197	Ach35197 Human end
C 588	16	1.4	415	4	AAI15791	AAi15791 Probe #57	C 661	16	1.4	498	5	AAI64358	AAi64358 DNA encod
C 589	16	1.4	415	4	ABA57973	Aba57973 Human foe	C 662	16	1.4	498	9	ACH27414	Ach27414 Human adu
C 590	16	1.4	415	4	AAI37571	AAi37571 Probe #62	C 663	16	1.4	501	12	ACH76613	Ach76613 Human gen
C 591	16	1.4	415	4	ABA27261	Aba27261 Probe #57	C 664	16	1.4	501	12	ACH77002	Ach77002 Human gen
C 592	16	1.4	415	4	AAK31695	Aak31695 Human bon	C 665	16	1.4	502	4	AAK84407	Aak84407 Human imm
C 593	16	1.4	415	4	AAK06047	Aak06047 Human bra	C 666	16	1.4	503	4	AAK84408	Aak84408 Human imm
C 594	16	1.4	415	4	ABS31379	Abs31379 Human liv	C 667	16	1.4	503	5	ADL62927	Adl62927 Human ova
C 595	16	1.4	415	6	ABS06451	Abs06451 Human gen	C 668	16	1.4	503	11	ACN91708	Acn91708 Breat ca
C 596	16	1.4	419	6	ABL77839	Ab177839 Human ova	C 669	16	1.4	507	9	ACH27997	Ach27997 Human adu
C 597	16	1.4	420	2	AAV09802	AAv09802 DNA encod	C 670	16	1.4	508	4	AAK88686	Aak88686 Human dig
C 598	16	1.4	422	4	AAI01959	AAi01959 Human rep	C 671	16	1.4	510	4	AAI17759	AAi17759 Probe #76
C 599	16	1.4	422	4	ABL97252	Ab197252 Human tes	C 672	16	1.4	510	4	ABA62707	Aba62707 Human foe
C 600	16	1.4	422	6	ABL69567	Ab169567 Prostate	C 673	16	1.4	510	4	AAI42711	AAi42711 Probe #11
C 601	16	1.4	425	3	AACT75575	Aac75575 Human ORF	C 674	16	1.4	510	4	ABA30008	Aba30008 Probe #84
C 602	16	1.4	425	6	ABN22023	Abn22023 Human ORF	C 675	16	1.4	510	4	AAK36910	Aak36910 Human bon
C 603	16	1.4	426	8	ABX50500	Abx50500 Bovine ES	C 676	16	1.4	510	4	AAK11086	Aak11086 Human bra
C 604	16	1.4	428	4	AAI16021	Aal16021 Human bre	C 677	16	1.4	510	4	ABS36582	Abs36582 Human liv

c 678	16	1.4	510	6	ABS10919	Abse10919 Human gen	c 751	16	1.4	619	4	AAK77975	Aak77975 Human imm
679	16	1.4	510	12	ACH71786	Ach71786 Human gen	c 752	16	1.4	619	6	ABNG4486	Abn64486 Human can
680	16	1.4	512	4	AAL16036	Aal16036 Human bre	c 753	16	1.4	622	13	ADO50944	Ado50944 Novel can
681	16	1.4	513	4	AAL14479	Aal14479 Human bre	c 754	16	1.4	627	4	AAL23362	Aal23362 Human bre
c 682	16	1.4	513	12	ADO57304	Ado57304 DNA encod	c 755	16	1.4	639	10	ADC91934	Adc91934 E. faeciu
683	16	1.4	515	4	AAL14454	Aal14454 Human bre	c 756	16	1.4	649	11	ACN84551	Acn84551 Breast ca
684	16	1.4	519	4	AAL144526	Aal14526 Human bre	c 757	16	1.4	655	13	ADQ56988	Adq56988 Novel can
c 685	16	1.4	519	4	AAC97016	Aac97016 Mycobacte	c 758	16	1.4	657	10	ACF70917	Acf70917 Photorhab
c 686	16	1.4	522	6	ABN61793	Abn61793 Human can	c 759	16	1.4	660	5	AAS84819	Aas84819 DNA encod
c 687	16	1.4	526	5	ABV58421	Abv58421 Human pro	c 760	16	1.4	665	11	ACN83289	Acn83289 Breast ca
c 688	16	1.4	526	8	ABZ56776	Abz56776 Aspergill	c 761	16	1.4	669	11	ACN83325	Acn83325 Breast ca
689	16	1.4	528	10	ABZ40611	Abz40611 N. gonorr	c 762	16	1.4	672	6	ABK35452	Abk35452 Human cdn
690	16	1.4	530	10	ADK12040	Adk12040 Breast ca	c 763	16	1.4	677	3	AAF12133	Aaf12133 Aspergill
c 691	16	1.4	530	11	ACN87894	Acn87894 Breast ca	c 764	16	1.4	678	5	AAS81591	Aas81591 DNA encod
c 692	16	1.4	531	3	AAAG9786	Aaa69786 Human ova	c 765	16	1.4	688	3	AAF12809	Aaf12809 Aspergill
693	16	1.4	531	6	ABN72680	Abn72680 Ovarian c	c 766	16	1.4	691	4	AAL12211	Aal12211 Human bre
694	16	1.4	531	9	ADA08845	Ada08845 Human ova	c 767	16	1.4	695	12	ADJ10627	Adj10627 Recombina
695	16	1.4	531	10	ADF08592	Adf08592 cDNA enco	c 768	16	1.4	704	4	AAL23271	Aal23271 Human bre
696	16	1.4	531	10	ADG46340	Adg46340 Human ova	c 769	16	1.4	707	4	AAH07453	Aah07453 Human cdn
c 697	16	1.4	534	3	AAZ53479	Aaz53479 Neisseria	c 770	16	1.4	714	10	ADK59676	Adk59676 Plant DNA
698	16	1.4	537	12	ACH79058	Ach79058 Human gen	c 771	16	1.4	717	4	AAL14500	Aal14500 Human bre
c 699	16	1.4	539	10	ADC76751	Adc76751 DNA homol	c 772	16	1.4	734	2	AAV53351	Aav53351 DNA encod
c 700	16	1.4	539	10	ADK58609	Adk58609 Plant DNA	c 773	16	1.4	738	4	AAF22430	Aaf22430 Human bre
c 701	16	1.4	546	6	ABN61916	Abn61916 Human can	c 774	16	1.4	743	8	ABV93918	Abv93918 Human col
702	16	1.4	548	4	ABA59532	Abas59532 Human foe	c 775	16	1.4	744	10	ADG32298	Adg32298 Mouse scf
703	16	1.4	548	4	AAI39394	Aai39394 Probe #80	c 776	16	1.4	750	13	ADR51430	Adr51430 Anti-biof
704	16	1.4	548	4	ABA28143	Aba28143 Probe #66	c 777	16	1.4	756	8	ACA38052	Aca38052 Prokaryot
705	16	1.4	548	4	AAK33671	Aak33671 Human bon	c 778	16	1.4	759	4	AAL23261	Aal23261 Human bre
706	16	1.4	548	4	AAK07806	Aak07806 Human bra	c 779	16	1.4	759	5	ADL36690	Adl36690 Human ova
707	16	1.4	548	4	ABS33486	Abs33486 Human liv	c 780	16	1.4	759	5	ADI71534	Adi71534 Human ova
708	16	1.4	548	6	ABS08568	Abs08568 Human gen	c 781	16	1.4	759	13	ADT48170	Adt48170 Bacterial
c 709	16	1.4	552	4	AAI16843	Aai16843 Probe #67	c 782	16	1.4	762	10	ADF66603	Adf66603 Corynebac
c 710	16	1.4	552	4	ABA60662	Abas60662 Human foe	c 783	16	1.4	762	12	ADQ31498	Adq31498 C. Glutam
c 711	16	1.4	552	4	AAI40551	Aai40551 Probe #92	c 784	16	1.4	764	10	ACF67089	Acf67089 Photorhab
c 712	16	1.4	552	4	ABA28758	Abas28758 Probe #72	c 785	16	1.4	771	11	ACN83214	Acn83214 Breast ca
c 713	16	1.4	552	4	AAK34835	Aak34835 Human bon	c 786	16	1.4	771	12	ADK16760	Adk16760 Nanoarcha
c 714	16	1.4	552	4	AAK08944	Aak08944 Human bra	c 787	16	1.4	775	4	AAL21981	Aal21981 Human bre
c 715	16	1.4	552	4	ABS34599	Abs34599 Human liv	c 788	16	1.4	776	4	AAL23389	Aal23389 Human bre
c 716	16	1.4	552	6	ABS09376	Abs09376 Human gen	c 789	16	1.4	777	3	AAZ53214	Aaz53214 Neisseria
c 717	16	1.4	552	8	ABZ52730	Abz52730 Aspergill	c 790	16	1.4	783	4	AHH04974	Aah04974 Human cdn
c 718	16	1.4	553	2	AAK21042	Aak21042 Polynucle	c 791	16	1.4	786	11	ACN83255	Acn83255 Breast ca
c 719	16	1.4	555	4	AAAL21094	Aal21094 Human bre	c 792	16	1.4	787	4	AAL13100	Aal13100 Human bre
720	16	1.4	560	4	AAH12171	Aah12171 Human cdn	c 793	16	1.4	792	11	ACN83267	Acn83267 Breast ca
c 721	16	1.4	561	13	ADQ91690	Adq91690 Polyketid	c 794	16	1.4	794	4	AAD08318	Aad08318 Human sec
c 722	16	1.4	568	13	ACN50618	Acn50618 Cotton ma	c 795	16	1.4	794	13	ADS62914	Ads62914 Bacterial
c 723	16	1.4	569	4	AAH12099	Aah12099 Human cdn	c 796	16	1.4	801	4	AAH47655	Aah47655 Clavulair
724	16	1.4	570	11	ACH98097	Ach98097 Klebsiell	c 797	16	1.4	801	4	AAD13054	Aad13054 Clavulair
725	16	1.4	576	11	ABD12592	Abd12592 Pseudomon	c 798	16	1.4	801	4	AAD11143	Aad11143 Clavulair
726	16	1.4	578	12	ACH76808	Ach76808 Human gen	c 799	16	1.4	801	12	ADM97770	Adm97770 Clavulair
c 727	16	1.4	584	6	ABX09754	Abx09754 M. incogn	c 800	16	1.4	804	3	AAZ53483	Aaz53483 Neisseria
c 728	16	1.4	588	10	ABZ40610	Abz40610 N. gonorr	c 801	16	1.4	804	3	AAZ53482	Aaz53482 Neisseria
729	16	1.4	588	12	ACH78715	Ach78715 Human gen	c 802	16	1.4	804	3	AAZ53480	Aaz53480 Neisseria
c 730	16	1.4	590	2	AAT93332	Aat93332 Exon 7 of	c 803	16	1.4	804	11	ACN84501	Acn84501 Breast ca
731	16	1.4	590	6	ABK44694	Abk44694 cDNA enco	c 804	16	1.4	804	11	ACN83283	Acn83283 Breast ca
c 732	16	1.4	591	4	AAK83676	Aak83676 Human imm	c 805	16	1.4	806	2	AAZ17555	Aaz17555 Human gen
c 733	16	1.4	591	4	AAK83675	Aak83675 Human imm	c 806	16	1.4	810	4	AAZ42027	Aaz42027 Genomic s
c 734	16	1.4	592	4	AAAL04677	Aal04677 Human rep	c 807	16	1.4	814	11	ACN84457	Acn84457 Breast ca
c 735	16	1.4	592	4	ABL97584	Ab197584 Human tes	c 808	16	1.4	819	11	ACN83282	Acn83282 Breast ca
736	16	1.4	592	12	ADI45583	Adi45583 Wheat iso	c 809	16	1.4	819	11	ACN83204	Acn83204 Breast ca
c 737	16	1.4	600	5	ABV54071	Abv54071 Human pro	c 810	16	1.4	823	11	ACN84456	Acn84456 Breast ca
738	16	1.4	601	13	ADS64402	Ads64402 Human tra	c 811	16	1.4	826	11	ACN84458	Acn84458 Breast ca
c 739	16	1.4	603	5	AAZ05498	Aaz05498 Mammalian	c 812	16	1.4	829	3	AAZ53426	Aaz53426 Arabidops
c 740	16	1.4	607	8	ABZ51619	Abz51619 Aspergill	c 813	16	1.4	830	11	ACN83229	Acn83229 Breast ca
c 741	16	1.4	608	13	ACN50555	Acn50555 Cotton ma	c 814	16	1.4	833	11	ACN86024	Acn86024 Breast ca
742	16	1.4	609	10	ABZ39246	Abz39246 N. gonorr	c 815	16	1.4	843	4	AAL23524	Aal23524 Human bre
c 743	16	1.4	609	10	ABZ39255	Abz39255 N. gonorr	c 816	16	1.4	848	11	ACN86015	Acn86015 Breast ca
744	16	1.4	612	2	AAT58248	Aat58248 5' end of	c 817	16	1.4	848	11	ACN84524	Acn84524 Breast ca
745	16	1.4	612	2	AAV10694	Aav10694 Human 0.6	c 818	16	1.4	850	11	ACN86009	Acn86009 Breast ca
746	16	1.4	612	3	AAZ299310	Aaz299310 5' end of	c 819	16	1.4	860	4	AAH99507	Aah99507 Human pro
c 747	16	1.4	612	3	AAZ53682	Aaz53682 Neisseria	c 820	16	1.4	863	4	AAH06479	Aah06479 Human cdn
c 748	16	1.4	612	3	AAZ53683	Aaz53683 Neisseria	c 821	16	1.4	867	11	ACH95953	Ach95953 Klebsiell
749	16	1.4	618	5	AAH88092	Aah88092 Peppermin	c 822	16	1.4	870	10	ADF02456	Adf02456 Bacterial
c 750	16	1.4	619	4	AAK77974	Aak77974 Human imm	c 823	16	1.4	870	11	ACN84393	Acn84393 Breast ca

824	16	1.4	872	4	AAH03828	Aah03828 Human cDN	897	16	1.4	1239	6	ABK98544	Abk98544 Botulinum
825	16	1.4	875	3	AAA02161	Aaa02161 Human col	898	16	1.4	1240	4	AAO09944	Aao09944 Human dru
826	16	1.4	876	11	ACN82390	Acn82390 Breast ca	899	16	1.4	1242	3	AAZ53204	Aaz53204 Neisseria
827	16	1.4	880	11	ACN85990	Acn85990 Breast ca	900	16	1.4	1242	3	AAZ53205	Aaz53205 Neisseria
828	16	1.4	884	11	ACN85890	Acn85890 Breast ca	901	16	1.4	1242	3	AAZ54590	Aaz54590 Botulism
829	16	1.4	886	11	ACN84567	Acn84567 Breast ca	902	16	1.4	1245	3	AAZ53203	Aaz53203 Neisseria
830	16	1.4	888	11	ACN83317	Acn83317 Breast ca	903	16	1.4	1245	4	ABA89407	Aba89407 Escherich
831	16	1.4	889	11	ACN85981	Acn85981 Breast ca	904	16	1.4	1248	4	AAH77987	Aah77987 Nucleotid
832	16	1.4	897	11	ACN83176	Acn83176 Breast ca	905	16	1.4	1257	13	AAZ49163	Aaz49163 Bacterial
833	16	1.4	897	11	ACN83176	Acn83176 Breast ca	906	16	1.4	1266	4	AAF33046	Aaf33046 Human sec
834	16	1.4	903	11	ACN83304	Acn83304 Breast ca	907	16	1.4	1266	8	ACA01423	ACA01423 C. glutam
835	16	1.4	915	6	ABN83168	Abn83168 Human sul	908	16	1.4	1275	12	ADN37670	Dengue vi
836	16	1.4	919	11	ACN83174	Acn83174 Breast ca	909	16	1.4	1275	12	ADN37683	Dengue vi
837	16	1.4	920	6	ABK14418	Abk14418 Letinula	910	16	1.4	1290	6	ABK74388	Abk74388 Bacillus
838	16	1.4	921	6	ABA03882	AbA03882 Human POL	911	16	1.4	1296	12	ADN37698	Dengue vi
839	16	1.4	921	8	ABX56478	Abx56478 cDNA enco	912	16	1.4	1302	5	AH68207	Aah68207 C glutami
840	16	1.4	921	10	AAZ58965	Aaz58965 Human sul	913	16	1.4	1310	11	ACN91780	Acn91780 Breast ca
841	16	1.4	922	4	AAZ19777	Aaz19777 Human ST	914	16	1.4	1313	5	AAZ79158	Aaz79158 DNA encod
842	16	1.4	926	13	ADS62272	Ads62272 Bacterial	915	16	1.4	1322	8	ABQ77371	Abq77371 C. glutam
843	16	1.4	926	13	ADS62588	Ads62588 Bacterial	916	16	1.4	1323	12	ADN37541	Dengue vi
844	16	1.4	929	11	ACN86000	Acn86000 Breast ca	917	16	1.4	1332	12	ADN37554	Dengue vi
845	16	1.4	942	4	AAZ19776	Aaz19776 Human ST	918	16	1.4	1332	12	ADN37680	Dengue vi
846	16	1.4	951	10	ADG323349	Adg323349 DNA encod	919	16	1.4	1332	12	ADN37697	Dengue vi
847	16	1.4	972	8	ACA49162	ACA49162 Prokaryot	920	16	1.4	1332	12	ADN37693	Dengue vi
848	16	1.4	975	10	ADB59083	Adb59083 Toxicity-	921	16	1.4	1332	12	ADN37676	Dengue vi
849	16	1.4	975	10	ABT42416	Abt42416 Toxicity	922	16	1.4	1332	12	ADN37685	Dengue vi
850	16	1.4	975	12	ADT430088	Adt430088 M. methyl	923	16	1.4	1332	12	ADN37705	Dengue vi
851	16	1.4	978	12	ADP72843	Adp72843 Renal tox	924	16	1.4	1332	12	ADN37694	Dengue vi
852	16	1.4	978	8	ACA51993	Aca51993 Prokaryot	925	16	1.4	1332	13	ADS51200	Ades1200 Bacterial
853	16	1.4	978	8	ACA51208	Aca51208 Prokaryot	926	16	1.4	1334	10	ADI02615	Adi02615 Human cDN
854	16	1.4	986	5	AAZ56791	Aaz56791 DNA encod	927	16	1.4	1338	10	ABZ38170	Abz38170 N. gonorr
855	16	1.4	995	11	ABD06602	Abd06602 Pseudomon	928	16	1.4	1338	12	ADN37660	Dengue vi
856	16	1.4	1004	4	AAZ19778	Aaz19778 Human ST	929	16	1.4	1338	12	ADN37682	Dengue vi
857	16	1.4	1009	11	ACN84568	Acn84568 Breast ca	930	16	1.4	1338	12	ADN37671	Dengue vi
858	16	1.4	1023	4	ABL20669	AbL20669 Drosophil	931	16	1.4	1338	12	ADN37684	Dengue vi
859	16	1.4	1038	10	ADE99758	Ade99758 Bacterial	932	16	1.4	1338	12	ADN37687	Dengue vi
860	16	1.4	1045	10	ACC79075	Acc79075 Human sec	933	16	1.4	1341	12	ADN37702	Dengue vi
861	16	1.4	1050	8	ACA52632	Aca52632 Prokaryot	934	16	1.4	1341	12	ADN37699	Dengue vi
862	16	1.4	1098	6	ABT07570	Abt07570 Human bre	935	16	1.4	1341	12	ADN37701	Dengue vi
863	16	1.4	1098	12	ADJ67913	Adj67913 T. mariti	936	16	1.4	1341	12	ADN37678	Dengue vi
864	16	1.4	1098	12	ADJ68125	Adj68125 T. mariti	937	16	1.4	1341	12	ADN37691	Dengue vi
865	16	1.4	1098	12	ADK01203	Adk01203 DNA polym	938	16	1.4	1341	12	ADN37674	Dengue vi
866	16	1.4	1098	12	ADJ79422	Adj79422 T. mariti	939	16	1.4	1341	12	ADN37661	Dengue vi
867	16	1.4	1098	12	ADJ79422	Adj79422 T. mariti	940	16	1.4	1344	11	ACH97627	Ach97627 Klebsiell
868	16	1.4	1098	12	ADJ84862	Adj84862 T. mariti	941	16	1.4	1344	12	ADN37569	Dengue vi
869	16	1.4	1098	12	ADM77650	Adm77650 DNA polym	942	16	1.4	1344	12	ADN37703	Dengue vi
870	16	1.4	1098	12	ADM66317	Adm66317 T. mariti	943	16	1.4	1347	12	ADN37668	Dengue vi
871	16	1.4	1098	12	ADO04370	Ado04370 T. mariti	944	16	1.4	1347	12	ADN37700	Dengue vi
872	16	1.4	1098	12	ADP82447	Adp82447 Thermotog	945	16	1.4	1347	12	ADN37675	Dengue vi
873	16	1.4	1107	13	ADT44895	Adt44895 Bacterial	946	16	1.4	1350	12	ADN37586	Dengue vi
874	16	1.4	1111	11	ACN84493	Acn84493 Breast ca	947	16	1.4	1365	11	ABD17927	Abd17927 Pseudomon
875	16	1.4	1116	3	AAZ52766	Aaz52766 Clavulari	948	16	1.4	1368	2	AAQ06226	Aaq06226 Amb a II
876	16	1.4	1116	4	AAZ03611	Aaz03611 Clavulari	949	16	1.4	1368	2	AAV13361	Aav13361 Ragweed p
877	16	1.4	1119	11	ACH97808	Ach97808 Klebsiell	950	16	1.4	1368	2	AAV37864	Aav37864 Allergen
878	16	1.4	1131	11	ACN84543	Acn84543 Breast ca	951	16	1.4	1368	6	ABA94120	Aba94120 Full leng
879	16	1.4	1136	5	ABV25109	Abv25109 Human pro	952	16	1.4	1368	6	ABA94141	Aba94141 Full leng
880	16	1.4	1143	13	ADS58715	Ads58715 Bacterial	953	16	1.4	1371	12	ADO30369	Ado30369 Mouse GPC
881	16	1.4	1146	5	AAH65662	Aah65662 C glutami	954	16	1.4	1374	5	AAI71573	Aai71573 Human zin
882	16	1.4	1153	11	ACN83218	Acn83218 Breast ca	955	16	1.4	1380	3	AAZ36697	Aaz36697 Candida a
883	16	1.4	1153	8	ABX63562	Abx63562 Human cDN	956	16	1.4	1380	5	AAZ75954	Aaz75954 DNA encod
884	16	1.4	1161	9	ABZ22611	Abz22611 Human NS4	957	16	1.4	1380	12	ADN37565	Dengue vi
885	16	1.4	1167	4	AAK85405	Aak85405 Human imm	958	16	1.4	1380	12	ADN37547	Dengue vi
886	16	1.4	1167	4	AAK85406	Aak85406 Human imm	959	16	1.4	1380	12	ADN37551	Dengue vi
887	16	1.4	1170	11	ACH95141	Ach95141 Klebsiell	960	16	1.4	1380	12	ADN37564	Dengue vi
888	16	1.4	1173	10	ADC93288	Adc93288 E. faeciu	961	16	1.4	1380	12	ADN37610	Dengue vi
889	16	1.4	1178	4	AAK90445	Aak90445 Human dig	962	16	1.4	1380	12	ADN37556	Dengue vi
890	16	1.4	1178	4	AAK90445	Aak90445 Human dig	963	16	1.4	1380	12	ADN37568	Dengue vi
891	16	1.4	1201	4	AAZ526939	Aaz526939 Human cDN	964	16	1.4	1382	4	ABL13839	AbL13839 Drosophil
892	16	1.4	1223	6	ADI16667	Adi16667 Human NOV	965	16	1.4	1383	4	ABL06178	Abl06178 Human rep
893	16	1.4	1223	12	ADN42321	Adn42321 Human cDN	966	16	1.4	1383	4	ABL98743	AbL98743 Human tes
894	16	1.4	1235	3	AAZ59128	Aaz59128 Human pan	967	16	1.4	1386	12	ADN37531	Dengue vi
895	16	1.4	1236	8	ACA37800	Aca37800 Prokaryot	968	16	1.4	1386	12	ADN37555	Dengue vi
896	16	1.4	1239	4	AAF67877	Aaf67877 Corynebac	969	16	1.4	1386	12	ADN37542	Dengue vi

c 970 16 1.4 1386 12 ADN37558
c 971 16 1.4 1386 12 ADN37553
c 972 16 1.4 1389 12 ADN37532
c 973 16 1.4 1389 12 ADN37549
c 974 16 1.4 1389 12 ADN37562
c 975 16 1.4 1389 12 ADN37545
c 976 16 1.4 1389 12 ADN37572
c 977 16 1.4 1389 12 ADN37570
c 978 16 1.4 1389 12 ADN37573
c 979 16 1.4 1390 6 ABZ15990
c 980 16 1.4 1392 12 ADN37574
c 981 16 1.4 1395 2 AAQ51243
c 982 16 1.4 1395 2 AAT32618
c 983 16 1.4 1395 12 ADN37546
c 984 16 1.4 1395 12 ADN37571
c 985 16 1.4 1395 12 ADN37539
c 986 16 1.4 1416 5 AAS81590
c 987 16 1.4 1419 8 ACA3154
c 988 16 1.4 1422 10 ACF72159
c 989 16 1.4 1425 11 ABD12383
c 990 16 1.4 1432 11 ADL65720
c 991 16 1.4 1432 11 ADL65464
c 992 16 1.4 1436 10 AAD60661
c 993 16 1.4 1436 11 ADJ26822
c 994 16 1.4 1452 12 ADN04702
c 995 16 1.4 1470 4 ABL26679
c 996 16 1.4 1470 11 ABD12658
c 997 16 1.4 1476 3 AAA08689
c 998 16 1.4 1488 4 AAS26845
c 999 16 1.4 1503 8 ACA27923
c1000 16 1.4 1506 2 AAX76394

ALIGNMENTS

RESULT 1
AAZ36102
ID AAZ36102 standard; DNA; 1181 BP.
XX
AC AAZ36102;
XX
DT 11-FEB-2000 (first entry)
XX
DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.
XX
KW Enterohemorrhagic Escherichia coli; EHEC; virulence factor;
KW enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.
XX
OS Escherichia coli.
XX
PN WO9955908-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-FR001000.
XX
PR 28-APR-1998; 98FR-00005329.
XX
PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Frechon DTM, Laure FC, Thierry D;
XX
DR WPI; 2000-013443/01.
XX
PT New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
PS Claim 1; Fig 2; 48pp; French.
XX
CC The present sequence is specific to enterohemorrhagic Escherichia coli
CC (EHEC). The sequence associated with the presence of virulence factors

CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
CC 68% homology with the virK gene which codes for virulence proteins of
CC Shigella flexneri. The present sequence is of plasmid origin. Fragments
CC of the present sequence are used, as probes and primers, for detection of
CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
CC animal samples, foods or the environment. The fragments are also useful
CC for epidemiological studies
XX
SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match 100.0%; Score 1181; DB 3; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1181; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGCAGGAGATGAAAAAGCCAAATATAAAATATGCCATCCAGCGCTCCAGCTG 60
DB 1 CTGCAGGAGATGAAAAAGCCAAATATAAAATATGCCATCCAGCGCTCCAGCTG 60

QY 61 AAAGTAGGCTGTTCCTGTCGGGTATTTAAATGCAATTTGACCGTCCCGTATTTAAACAATG 120
DB 61 AAAGTAGGCTGTTCCTGTCGGGTATTTAAATGCAATTTGACCGTCCCGTATTTAAACAATG 120

QY 121 TGAATAATTAATCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAAGAGATCC 180
DB 121 TGAATAATTAATCCGTTACCGGAAAACCGCTGAACAAAATTCGGGCTGAAAAAGAGATCC 180

QY 181 GCGGTTATCTGTTGCAATTTCCCTTAGCTAGCTAGCAGAGACACAATGATCTGTGCGG 240
DB 181 GCGGTTATCTGTTGCAATTTCCCTTAGCTAGCTAGCAGAGACACAATGATCTGTGCGG 240

QY 241 TTCTGTTAATATCAAAACCGGTACTCAATATCTTCTGCGGCTGGCTGCCATCATCCGGA 300
DB 241 TTCTGTTAATATCAAAACCGGTACTCAATATCTTCTGCGGCTGGCTGCCATCATCCGGA 300

QY 301 AGCGTTCCGGTCGGGATAAAAAATCGCGCAGTCGGCGGTCATGACAGACACATCCCCCA 360
DB 301 AGCGTTCCGGTCGGGATAAAAAATCGCGCAGTCGGCGGTCATGACAGACACATCCCCCA 360

QY 361 CGGGTAACAGCGTCCTGTCACATCTTCTGTAATGACATCAGGGATCCGCGCTCTCAC 420
DB 361 CGGGTAACAGCGTCCTGTCACATCTTCTGTAATGACATCAGGGATCCGCGCTCTCAC 420

QY 421 TGGCGATAACGGGCACGCGGAGACTGACGCTTTACGCCAGTACATACCAACGCTTCAT 480
DB 421 TGGCGATAACGGGCACGCGGAGACTGACGCTTTACGCCAGTACATACCAACGCTTCAT 480

QY 481 TTTCCGAAGCATGACCAACCACTGCGCATCCGGTAGACCGGTAAAGTGGGAAAGGG 540
DB 481 TTTCCGAAGCATGACCAACCACTGCGCATCCGGTAGACCGGTAAAGTGGGAAAGGG 540

QY 541 CACCTGCCATTAACACATCTCCGCTCATTTCCAGGTGTTCTGCTGACGACGAGCTG 600
DB 541 CACCTGCCATTAACACATCTCCGCTCATTTCCAGGTGTTCTGCTGACGACGAGCTG 600

QY 601 CTTTCGTATTTTCACGCCCGCGGCCCAACGAGCCAGCGAAATGATTTCCCTTCATCT 660
DB 601 CTTTCGTATTTTCACGCCCGCGGCCCAACGAGCCAGCGAAATGATTTCCCTTCATCT 660

QY 661 TCAGCTGATACATAACGACGATTAATTCATGCTCTTTTCGGGACGTAGCATCCCCA 720
DB 661 TCAGCTGATACATAACGACGATTAATTCATGCTCTTTTCGGGACGTAGCATCCCCA 720

QY 721 CTTGAACGATAAGCGGAACAATTTGCTGCTGATGACGCCAGGGTGGATATGAGGGGTA 780
DB 721 CTTGAACGATAAGCGGAACAATTTGCTGCTGATGACGCCAGGGTGGATATGAGGGGTA 780

QY 781 ACGGTCGATGCGCTTCATTTGCAATCGCGGCGAGTCGAAACCCCGGTGGAATTAACCGTTA 840
DB 781 ACGGTCGATGCGCTTCATTTGCAATCGCGGCGAGTCGAAACCCCGGTGGAATTAACCGTTA 840

QY 841 CCGGTGCTCTGACACCTTCGGCCATCAGATGCGCCATCATGGGTGAGTAGGACACAA 900
DB 841 CCGGTGCTCTGACACCTTCGGCCATCAGATGCGCCATCATGGGTGAGTAGGACACAA 900

Qy	901	TGAAATCACACAGATAAATTCAGGGAAACGTTCTGGTCTTACGGGTGATGATAGCTTTTTT	960
Db	901	TGAAATCACACAGATAAATTCAGGGAAACGTTCTGGTCTTACGGGTGATGATAGCTTTTTT	960
Qy	961	GTCTGACAAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCCTGCTATATTACTGT	1020
Db	961	GTCTGACAAATAGTGAAGCGGTGACAGCATATCAGACGGCTCAGTCCCTGCTATATTACTGT	1020
Qy	1021	CATGGCCACTATGTCAGAGATGACAGATCAGGTTTAAATTCGCCGATTAATCCGTCGAAGTC	1080
Db	1021	CATGGCCACTATGTCAGAGATGACAGATCAGGTTTAAATTCGCCGATTAATCCGTCGAAGTC	1080
Qy	1081	TGAGATGGAAGGAAGGTGAAGGCTGTTTCTGAAAGGAATAAAAGTCACATCATGCCCTC	1140
Db	1081	TGAGATGGAAGGAAGGTGAAGGCTGTTTCTGAAAGGAATAAAAGTCACATCATGCCCTC	1140
Qy	1141	TTTTTCTGGCTTCGGAGCAATTTTACTTTTTTCTCTGCAG	1181
Db	1141	TTTTTCTGGCTTCGGAGCAATTTTACTTTTTTCTCTGCAG	1181

RESULT 2
AAZ36126
ID AAZ36126 standard; DNA: 31 BP.

DT 11-FEB-2000 (first entry)

DE Primer derived from a nucleic acid sequence specific to EHEC.

Enterohemorrhagic *Escherichia coli*; EHEC; katP gene; *E. coli* O157:H7;
I991; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
PCR primer; probe; ss.

OS Synthetic.
OS *Escherichia coli*.

PN WO9955908-A2.

04-NOV-1999.

PF 27-APR-1999; 99WO-FR001000.

PR 28-APR-1998; 98FR-00005329.

PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PI Frechon DTM, Laure FC, Thierry D;

DR WPI; 2000-013443/01.

PT New nucleic acid containing sequences specific to enterohemorrhagic
PT *Escherichia coli*, particularly serotype O157:H7, used for detecting these
PT bacteria in food.

PS Claim 5; Page 27; 48pp; French.

AAZ36103-27 represent fragments derived from nucleic acid sequences specific to enterohemorrhagic *Escherichia coli* (EHEC). The fragments are derived from two sequences. The first (AAZ36101) is 99.9% homologous to the katP gene of *E. coli* O157:H7 (nucleotides 407-1489 of AAZ36101), and 95.8% homologous with 1991 of *E. coli* (nucleotides 1-406 of AAZ36102).

The second sequence (AA236102) is associated with the presence of virulence factors enterohemolysins (ehly) and intimin (eae). Nucleotides 237-570 of AA236102 also have 68% homology with the virk gene which codes for virulence proteins of *Shigella flexneri*. Both sequences are of plasmid origin. The fragments are used as PCR primers and probes for the detection of *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC), in human or animal samples, foods or the environment. The fragments are also useful for epidemiological studies

SQ Sequence 31 BP; 9 A; 6 C; 10 G; 6 T; 0 U; 0 Other;

Query Match 2.6%; Score 31; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels

Qy 970 TACTGAAGCGGTGACAGCATATCAGACGGCT 1000
|||||
Db 1 TACTGAAGCGGTGACAGCATATCAGACGGCT 31

RESULT 3
AAZ36125
ID AAZ36125 standard; DNA; 31 BP.

DT 11-FEB-2000 (first entry)

DE Primer derived from a nucleic acid sequence specific to EHEC.

AA Enterohemorrhagic *Escherichia coli*; EHEC; katP gene; *E. coli* O157:H7;
KW I991; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.

OS Synthetic.
OS *Escherichia coli*.

PN WO9955908-A2.

XX
PD
04-NOV-1999.

27-APR-1999; 99WO-FR001000.

PR 28-APR-1998; 98FR-00005329.

PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PI Frechon DTM, Laure FC, Thierry D;

DR WPI; 2000-013443/01.

PT New nucleic acid containing sequences specific to enterohemorrhagic
PT *Escherichia coli*, particularly serotype O157:H7, used for detecting these
PT bacteria in food.

PS Claim 5; Page 27; 48pp; French.

CC AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic *Escherichia coli* (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the *katP* gene of *E. coli* O157:H7 (nucleotides 407-1489 of AAZ36101), and

95.8% homologous with IS91 of *E. coli* (nucleotides 1-406 of AAZ36102). The second sequence (AAZ36102) is associated with the presence of virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 of AAZ36102 also have 68% homology with the *virK* gene which codes for virulence proteins of *Shigella flexneri*. Both sequences are of plasmid origin. The fragments are used as PCR primers and probes for detection of *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC), in human or animal samples, foods or the environment. The fragments are also useful for epidemiological studies

Sequence 31 BP; 4 A; 4 C; 10 G; 13 T; 0 U; 0 Other;

Query Match 2.6%; Score 31; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 31; Conservative 0; Mismatches 0; Indels

Qy 928 ACGTTCTGGTCTTACGGGTGATGATAGGTTTT 958
|||||
Db 1 ACGTTCTGGTCTTACGGGTGATGATAGGTTTT 31
|||||

```
RESULT 4
ACF69772/c
ID ACF69772 standard; DNA; 2136 BP.
XX AC ACF69772;
XX AC
XX 20-NOV-2003 (first entry)
XX DT
XX DE Photorhabdus luminescens nucleotide sequence #8239.
XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
XX detection; food; gene expression; plant; animal; microorganism; toxin;
XX antibiotic; biopesticide; virulence factor; disease model; plague;
XX whooping cough; gene; ds.
XX OS
XX Photorhabdus luminescens.
XX WO200294867-A2.
XX PN
XX 28-NOV-2002.
XX PD
XX PF 07-FEB-2002; 2002WO-IB003040.
XX PR 07-FEB-2001; 2001FR-00001659.
XX PA (INSP ) INST PASTEUR.
XX PA (CNRS ) CNRS CENT NAT RECH SCI.
XX PI Duclaud E, Taourit S, Glaser P, Prangeul L, Kunst F, Danchin A;
PI Buchrieser C;
XX WI; 2003-148459/14.
XX DR
XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
XX useful e.g. as therapeutic antimicrobials and agricultural pesticides.
XX PS Claim 2; SEQ ID NO 8239; 1205pp; French.
XX CC The invention relates to the isolation of genes and their encoded
XX proteins from Photorhabdus luminescens. The isolated sequences are
XX sources of probes and primers for detecting the genome of P. luminescens
XX and related species; to study polymorphisms; for gene analysis and for
XX detection/amplification of the genes. Antibodies (Ab) raised against the
XX polypeptides encoded by the genes are used for detection/identification
XX of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
XX carry a gene-containing vector are used to select compounds that
XX modulate, regulate, induce or inhibit expression of the genes in plants,
XX animals or microorganisms other than P. luminescens and are able to alter
XX response or sensitivity to toxins and antibiotics produced by P.
XX luminescens. Cells transformed to express the genes are useful for
XX recombinant production of the proteins, particularly toxins and
XX antibacterials useful as insecticides, bactericides and fungicides. The
XX genes, proteins, vectors containing the genes and Ab are also useful
XX therapeutically (to treat microbial infection by bacteria or fungi that
XX are sensitive to P. luminescens-encoded toxins or antibiotics) and as
XX biopesticides. Other uses of the genes and the proteins are as virulence
XX factors and for identifying targets of human diseases for which P.
XX luminescens is a model (particularly plague and whooping cough). This
XX sequence represents one of the isolated P. luminescens genes
XX SQ Sequence 2136 BP; 424 A; 569 C; 605 G; 538 T; 0 U; 0 Other;
Query Match 2.0%; Score 24; DB 10; Length 2136;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 AGCCAAATAAAAAATTGCCCATC 43
| | | | | | | | | | | | | | | | | | | | | |
Db 276 AGCCAAATAAAAAATTGCCCATC 253
| | | | | | | | | | | | | | | | | | | | | |
RESULT 5
ACF67367_27
Continuation (2 of 7) of ACF65386 from base 100001 (Photorhabdus luminescens nucleotide
```

```
Continuation (28 of 57) of ACF67367 from base 2700001 (Photorhabdus luminescens nucleot
WP Sequence split into 57 fragments LOCUS ACF67367 Accession Acf67367
WP Fragment Name Begin End
WP ACF67367_00 1 110000
WP ACF67367_01 100001 210000
WP ACF67367_02 200001 310000
WP ACF67367_03 300001 410000
WP ACF67367_04 400001 510000
WP ACF67367_05 500001 610000
WP ACF67367_06 600001 710000
WP ACF67367_07 700001 810000
WP ACF67367_08 800001 910000
WP ACF67367_09 900001 1010000
WP ACF67367_10 1000001 1110000
WP ACF67367_11 1100001 1210000
WP ACF67367_12 1200001 1310000
WP ACF67367_13 1300001 1410000
WP ACF67367_14 1400001 1510000
WP ACF67367_15 1500001 1610000
WP ACF67367_16 1600001 1710000
WP ACF67367_17 1700001 1810000
WP ACF67367_18 1800001 1910000
WP ACF67367_19 1900001 2010000
WP ACF67367_20 2000001 2110000
WP ACF67367_21 2100001 2210000
WP ACF67367_22 2200001 2310000
WP ACF67367_23 2300001 2410000
WP ACF67367_24 2400001 2510000
WP ACF67367_25 2500001 2610000
WP ACF67367_26 2600001 2710000
WP ACF67367_27 2700001 2810000
WP ACF67367_28 2800001 2910000
WP ACF67367_29 2900001 3010000
WP ACF67367_30 3000001 3110000
WP ACF67367_31 3100001 3210000
WP ACF67367_32 3200001 3310000
WP ACF67367_33 3300001 3410000
WP ACF67367_34 3400001 3510000
WP ACF67367_35 3500001 3610000
WP ACF67367_36 3600001 3710000
WP ACF67367_37 3700001 3810000
WP ACF67367_38 3800001 3910000
WP ACF67367_39 3900001 4010000
WP ACF67367_40 4000001 4110000
WP ACF67367_41 4100001 4210000
WP ACF67367_42 4200001 4310000
WP ACF67367_43 4300001 4410000
WP ACF67367_44 4400001 4510000
WP ACF67367_45 4500001 4610000
WP ACF67367_46 4600001 4710000
WP ACF67367_47 4700001 4810000
WP ACF67367_48 4800001 4910000
WP ACF67367_49 4900001 5010000
WP ACF67367_50 5000001 5110000
WP ACF67367_51 5100001 5210000
WP ACF67367_52 5200001 5310000
WP ACF67367_53 5300001 5410000
WP ACF67367_54 5400001 5510000
WP ACF67367_55 5500001 5610000
WP ACF67367_56 5600001 5648894
Query Match 2.0%; Score 24; DB 10; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 AGCCAAATAAAAAATTGCCCATC 43
| | | | | | | | | | | | | | | | | | | | | |
Db 76571 AGCCAAATAAAAAATTGCCCATC 76594
| | | | | | | | | | | | | | | | | | | | | |
RESULT 6
ACF65386_1/c
Continuation (2 of 7) of ACF65386 from base 100001 (Photorhabdus luminescens nucleotide
```

WP Sequence split into 7 fragments LOCUS ACF65386 Accession ACF65386
WP Fragment Name Begin End
WP ACF65386_0 1 110000
WP ACF65386_1 100001 210000
WP ACF65386_2 200001 310000
WP ACF65386_3 300001 410000
WP ACF65386_4 400001 510000
WP ACF65386_5 500001 610000
WP ACF65386_6 600001 700779

Query Match 2.0%; Score 24; DB 10; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.16; Mismatches 0; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 AGCCAAATAAAAAATGGCCATC 43
ID AAZ36122 standard; DNA; 22 BP.
XX
AC AAZ36122;
XX
DT 11-FEB-2000 (first entry)
XX
DE Primer derived from a nucleic acid sequence specific to EHEC.
XX
KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX
OS Synthetic.
OS Escherichia coli.
XX
PN W09955908-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-FR001000.
XX
PR 28-APR-1998; 98FR-00005329.
XX
PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Frechon DTM, Laure FC, Thierry D;
XX
DR WPI; 2000-013443/01.
XX
PT New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
PS Claim 5; Page 27; 48pp; French.
XX
XX AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of
CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies
SQ Sequence 22 BP; 4 A; 11 C; 1 G; 6 T; 0 U; 0 Other;

Query Match 1.9%; Score 22; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 863 CATCAGATGCCCATCATGGGT 884
ID 22 CATCAGATGCCCATCATGGGT 1

RESULT 9
AAZ36121
ID AAZ36121 standard; DNA; 22 BP.

Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1078 GTCTGAGGATGGAAGGAAGGTG 1099
DB 22 GTCTGAGGATGGAAGGAAGGTG 1

RESULT 8
AAZ36124/c
ID AAZ36124 standard; DNA; 22 BP.
XX
AC AAZ36124;
XX
DT 11-FEB-2000 (first entry)
XX
DE Primer derived from a nucleic acid sequence specific to EHEC.
XX
KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX
OS Synthetic.
OS Escherichia coli.
XX
PN W09955908-A2.
XX
PD 04-NOV-1999.
XX
PF 27-APR-1999; 99WO-FR001000.
XX
PR 28-APR-1998; 98FR-00005329.
XX
PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
XX
PI Frechon DTM, Laure FC, Thierry D;
XX
DR WPI; 2000-013443/01.
XX
PT New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX
PS Claim 5; Page 27; 48pp; French.
XX
XX AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of
CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies
SQ Sequence 22 BP; 5 A; 6 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 1.9%; Score 22; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 863 CATCAGATGCCCATCATGGGT 884
DB 22 CATCAGATGCCCATCATGGGT 1

RESULT 9
AAZ36121
ID AAZ36121 standard; DNA; 22 BP.

PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-102386/09.
DR
XX
PT Purified nucleic acid molecules, useful for genome mapping, gene
PT identification and analysis, cattle breeding or preparation of constructs
PT for cattle gene expression and genetically improved cattle.
XX
XX Claim 2; SEQ ID NO 1524; 38pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 5912 nucleotide sequences,
CC appearing as ABX50072-ABX55983, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 5912 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the 5912
CC bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present
CC sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137160
XX
SQ Sequence 341 BP; 83 A; 86 G; 97 G; 73 T; 0 U; 2 Other;

Query Match 1.8%; Score 21; DB 8; Length 341;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1077 AGTCTGAGGATGGAAGGAGG 1097
Db 245 AGTCTGAGGATGGAAGGAGG 265
|||||

RESULT 12
AAZ36123
ID AAZ36123 standard; DNA; 20 BP.
XX
AC AAZ36123;
XX
XX 11-FEB-2000 (first entry)
DT
DT
DE Primer derived from a nucleic acid sequence specific to EHEC.
XX
XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW I591; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX
XX Synthetic.
OS Escherichia coli.
OS
XX WO955908-A2.
PN
XX
XX 04-NOV-1999.
PD

XX 27-APR-1999; 99WO-FR001000.
XX 28-APR-1998; 98FR-00005329.
XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.
XX Frechon DTM, Laure FC, Thierry D;
PI WPI; 2000-013443/01.
XX
XX New nucleic acid containing sequences specific to enterohemorrhagic
XX Escherichia coli, particularly serotype O157:H7, used for detecting these
XX bacteria in food.
XX
XX Claim 5; Page 27; 48pp; French.
XX
XX AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with I591 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of
CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies
XX
SQ Sequence 20 BP; 3 A; 9 C; 5 G; 3 T; 0 U; 0 Other;

Query Match 1.7%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ATCCGAGCGCGTCCAGCTG 60
Db 1 ATCCGAGCGCGTCCAGCTG 20
|||||

RESULT 13
ABQ70353/c
ID ABQ70353 standard; DNA; 1077 BP.
XX
XX ABQ70353;
XX
DT 29-AUG-2003 (revised)
DT 29-AUG-2002 (first entry)
XX
XX Listeria monocytogenes 4b contig DNA sequence #295.
XX Antibacterial; Listeria; food contamination; mutational analysis;
XX infection; ds.
XX
XX Listeria monocytogenes ATCC 19115.
XX
XX WO200228891-A2.
PN
XX 11-APR-2002.
PD
XX
XX 04-OCT-2001; 2001WO-FR003061.
PF
XX 04-OCT-2000; 2000FR-00012697.
PR
XX (INSP) INST PASTEUR.
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
XX Kunst F, Glaser P;
PI WPI; 2002-332479/37.
XX
XX

PT New genomic sequences from *Listeria* species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators.
XX
XX Claim 14; SEQ ID NO 3166; 180pp; French.
XX
CC The present invention relates to nucleic acid sequences (ABQ67188-
CC ABQ71212) from *Listeria* sp. The sequences are useful as probes and
CC primers for identification and/or detection of *Listeria* (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of gene
CC expression. Proteins encoded by the nucleic acid sequences can be used to
CC screen for compounds that modulate gene expression, replication and
CC pathogenicity of *Listeria* (potential therapeutic agents), also for
CC treating infections by *Listeria*, and are useful as immunogens in anti-
CC *Listeria* vaccines. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. (Updated
CC on 29-AUG-2003 to standardise OS field)
XX
XX
SQ Sequence 1077 BP; 356 A; 194 C; 233 G; 294 T; 0 U; 0 Other;
Query Match 1.7%; Score 20; DB 6; Length 1077;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 683 CATAAATTCATGTCCTTTT 702
|||||
DB 773 CATAAATTCATGTCCTTTT 754
RESULT 14
AAAX38295/c
ID AAX38295 standard; DNA; 3589 BP.
XX
AC AAX38295;
XX
DT 16-JUN-1999 (first entry)
XX
DE Staphylococcus carnosus DNA containing the secA gene.
XX
KW Microorganism inhibitor; antisense; nuclease resistant; treatment;
KW ribonucleotide reductase; secA gene; pathological condition;
KW antimicrobial agent; crop protection; ss.
XX
OS Staphylococcus carnosus.
XX
PN WO9902673-A2.
XX
PD 21-JAN-1999.
XX
PF 10-JUL-1998; 98WO-CA000666.
XX
PR 10-JUL-1997; 97US-0052160P.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
PI Wright JA, Young AH, Dugourd D;
XX
DR WPI; 1999-120874/10.
XX
XX
PT New oligonucleotides complementary to RR or SecA genes - useful to
PT inhibit growth of microorganisms.
XX
PS Disclosure; Fig 9; 103pp; English.
XX
CC This invention describes novel antisense oligonucleotides (AAX38301-
CC X38552) which are nuclease resistant, and comprises about 3-50
CC nucleotides complementary to the ribonucleotide reductase gene or the
CC secA gene of a microorganism. The antisense oligonucleotides are used to
CC treat mammalian pathological conditions mediated by microorganisms. The
CC oligonucleotides are particularly useful as antimicrobial agents in crop
CC protection. This DNA sequence contains the Staphylococcus carnosus secA
CC gene

XX
SQ Sequence 3589 BP; 1308 A; 565 C; 725 G; 990 T; 0 U; 1 Other;
Query Match 1.7%; Score 20; DB 2; Length 3589;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 653 TTCATCTTCAGCTGATACA 672
|||||
DB 3301 TTCATCTTCAGCTGATACA 3282
RESULT 15
ABL14236
ID ABL14236 standard; cDNA; 5402 BP.
XX
AC ABL14236;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37190.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-856860/75.
DR P-PSDB; ABB70133.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 37190; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5402 BP; 1296 A; 1265 C; 1152 G; 1689 T; 0 U; 0 Other;
Query Match 1.7%; Score 20; DB 4; Length 5402;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 886 AGATAGGCACCAATGAAA 905
|||||
DB 3041 AGATAGGCACCAATGAAA 3060
Search completed: June 4, 2005, 15:46:32
Job time : 769 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 15:20:37 ; Search time 238 Seconds
(without alignments)

(without alignment)
8119.508 Million cell updates/sec

Title: US-09-674-277-2

Perfect score: 1181

Sequence: 1 ctgcaggagatggaaaaaa.....tttactttttctctgcag 1181

Scoring table:

Gapop_60.0 , Gapext 60.0

.....T.....T.....T.....

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

```
1: /cqn2 6/ptodata/1/ina/5A COMB.seq:*
```

2: /cqn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cqn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:*

```
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		ID	Description	
		Match	Length			
C 1	20	1.7	3589	4	US-09-112-580-9	
C 2	19	1.6	267	4	US-09-248-786A-12412	
C 3	19	1.6	601	4	US-09-949-016-161314	
C 4	19	1.6	2680	4	US-09-949-016-4520	
C 5	19	1.6	2982	4	US-09-620-312D-307	
C 6	7	19	1.6	5828	4	US-09-949-016-15210
C 7	19	1.6	45571	4	US-09-949-016-16262	
C 8	19	1.6	65888	4	US-09-949-016-11942	
C 9	19	1.6	66989	4	US-09-949-016-16063	
C 10	19	1.6	68702	4	US-09-949-016-16328	
C 11	19	1.6	135171	4	US-09-949-016-15617	
C 12	18	1.5	531	4	US-09-270-767-11620	
C 13	18	1.5	601	4	US-09-949-016-55105	
C 14	18	1.5	601	4	US-09-949-016-55106	
C 15	18	1.5	601	4	US-09-949-016-76772	
C 16	18	1.5	601	4	US-09-949-016-117483	
C 17	18	1.5	601	4	US-09-949-016-117484	
C 18	18	1.5	601	4	US-09-949-016-201194	
C 19	18	1.5	954	4	US-09-134-000C-1282	
C 20	18	1.5	1001	3	US-09-641-638-401	
C 21	18	1.5	1001	4	US-10-170-097-401	
C 22	18	1.5	1154	3	US-09-347-803-15	
C 23	18	1.5	2154	4	US-09-134-000C-2428	
C 24	18	1.5	2281	3	US-09-392-772-9	
C 25	18	1.5	2894	4	US-09-620-312D-95	
C 26	18	1.5	3606	4	US-10-164-595-37	
C 27	18	1.5	19866	4	US-09-949-016-16729	

c 101	17	1.4	25512	4	US-09-949-016-15887	Sequence 15887, A	174	16	1.4	576	4	US-09-252-991A-11196	Sequence 11196, A
c 102	17	1.4	31197	4	US-09-949-016-12963	Sequence 12963, A	c 175	16	1.4	588	4	US-09-902-8540-8564	Sequence 8564, Ap
c 103	17	1.4	34063	3	US-09-453-702B-96	Sequence 96, Appl	c 176	16	1.4	590	3	US-08-906-156A-19	Sequence 19, Appl
c 104	17	1.4	34185	3	US-09-545-481-3	Sequence 3, Appli	c 177	16	1.4	601	4	US-09-949-016-23631	Sequence 23631, A
c 105	17	1.4	34199	4	US-09-902-940-1255	Sequence 1255, Ap	c 178	16	1.4	601	4	US-09-949-016-26552	Sequence 26552, A
c 106	17	1.4	40037	4	US-09-949-016-12715	Sequence 12715, A	c 179	16	1.4	601	4	US-09-949-016-26903	Sequence 26903, A
c 107	17	1.4	45175	3	US-09-453-702B-116	Sequence 116, App	c 180	16	1.4	601	4	US-09-949-016-26904	Sequence 26904, A
c 108	17	1.4	51711	4	US-09-949-016-12559	Sequence 12559, A	c 181	16	1.4	601	4	US-09-949-016-31260	Sequence 31260, A
c 109	17	1.4	57103	4	US-09-949-016-13445	Sequence 13445, A	c 182	16	1.4	601	4	US-09-949-016-32040	Sequence 32040, A
c 110	17	1.4	70563	4	US-09-949-016-16743	Sequence 16743, A	c 183	16	1.4	601	4	US-09-949-016-32041	Sequence 32041, A
c 111	17	1.4	74881	4	US-09-949-016-15545	Sequence 15545, A	c 184	16	1.4	601	4	US-09-949-016-32953	Sequence 32953, A
c 112	17	1.4	74914	4	US-09-949-016-12286	Sequence 12286, A	c 185	16	1.4	601	4	US-09-949-016-43082	Sequence 43082, A
c 113	17	1.4	81001	3	US-09-750-580-1	Sequence 1, Appli	c 186	16	1.4	601	4	US-09-949-016-43289	Sequence 43289, A
c 114	17	1.4	84558	4	US-09-949-016-15752	Sequence 15752, A	c 187	16	1.4	601	4	US-09-949-016-43518	Sequence 43518, A
c 115	17	1.4	103934	4	US-09-949-016-14433	Sequence 14433, A	c 188	16	1.4	601	4	US-09-949-016-43747	Sequence 43747, A
c 116	17	1.4	112132	3	US-09-741-150-3	Sequence 3, Appli	c 189	16	1.4	601	4	US-09-949-016-43815	Sequence 43815, A
c 117	17	1.4	112132	4	US-10-160-187-3	Sequence 3, Appli	c 190	16	1.4	601	4	US-09-949-016-43860	Sequence 43860, A
c 118	17	1.4	144158	4	US-09-949-016-11755	Sequence 11755, A	c 191	16	1.4	601	4	US-09-949-016-44289	Sequence 44289, A
c 119	17	1.4	144158	4	US-09-949-016-12936	Sequence 12936, A	c 192	16	1.4	601	4	US-09-949-016-45803	Sequence 45803, A
c 120	17	1.4	152393	4	US-09-949-016-14514	Sequence 14514, A	c 193	16	1.4	601	4	US-09-949-016-45804	Sequence 45804, A
c 121	17	1.4	152393	4	US-09-949-016-14515	Sequence 14515, A	c 194	16	1.4	601	4	US-09-949-016-48244	Sequence 48244, A
c 122	17	1.4	156894	4	US-09-949-016-12765	Sequence 12765, A	c 195	16	1.4	601	4	US-09-949-016-52027	Sequence 52027, A
c 123	17	1.4	156894	4	US-09-949-016-12766	Sequence 12766, A	c 196	16	1.4	601	4	US-09-949-016-52432	Sequence 52432, A
c 124	17	1.4	156895	4	US-09-949-016-16957	Sequence 16957, A	c 197	16	1.4	601	4	US-09-949-016-52477	Sequence 52477, A
c 125	17	1.4	156895	4	US-09-949-016-16958	Sequence 16958, A	c 198	16	1.4	601	4	US-09-949-016-69229	Sequence 69229, A
c 126	17	1.4	164061	4	US-09-949-016-16959	Sequence 16959, A	c 199	16	1.4	601	4	US-09-949-016-69230	Sequence 69230, A
c 127	17	1.4	164061	4	US-09-949-016-17422	Sequence 17422, A	c 200	16	1.4	601	4	US-09-949-016-82022	Sequence 82022, A
c 128	17	1.4	187916	4	US-09-949-016-12980	Sequence 12980, A	c 201	16	1.4	601	4	US-09-949-016-82023	Sequence 82023, A
c 129	17	1.4	192302	4	US-09-949-016-15270	Sequence 15270, A	c 202	16	1.4	601	4	US-09-949-016-82024	Sequence 82024, A
c 130	17	1.4	192302	4	US-09-949-016-15270	Sequence 15270, A	c 203	16	1.4	601	4	US-09-949-016-82025	Sequence 82025, A
c 131	17	1.4	205044	4	US-09-949-016-15851	Sequence 15851, A	c 204	16	1.4	601	4	US-09-949-016-84397	Sequence 84397, A
c 132	17	1.4	205044	4	US-09-949-016-15852	Sequence 15852, A	c 205	16	1.4	601	4	US-09-949-016-84397	Sequence 84397, A
c 133	17	1.4	205044	4	US-09-949-016-15853	Sequence 15853, A	c 206	16	1.4	601	4	US-09-949-016-87741	Sequence 87741, A
c 134	17	1.4	223471	4	US-09-949-016-12387	Sequence 12387, A	c 207	16	1.4	601	4	US-09-949-016-93138	Sequence 93138, A
c 135	17	1.4	223471	4	US-09-949-016-12724	Sequence 12724, A	c 208	16	1.4	601	4	US-09-949-016-108354	Sequence 108354, A
c 136	17	1.4	223471	4	US-09-949-016-12725	Sequence 12725, A	c 209	16	1.4	601	4	US-09-949-016-113807	Sequence 113807, A
c 137	17	1.4	260247	4	US-09-949-016-13358	Sequence 13358, A	c 210	16	1.4	601	4	US-09-949-016-113808	Sequence 113808, A
c 138	17	1.4	276237	4	US-09-949-016-17504	Sequence 17504, A	c 211	16	1.4	601	4	US-09-949-016-113809	Sequence 113809, A
c 139	17	1.4	323820	4	US-09-949-016-14139	Sequence 14139, A	c 212	16	1.4	601	4	US-09-949-016-113975	Sequence 113975, A
c 140	17	1.4	346112	4	US-09-949-016-13165	Sequence 13165, A	c 213	16	1.4	601	4	US-09-949-016-113976	Sequence 113976, A
c 141	17	1.4	363032	4	US-09-949-016-12415	Sequence 12415, A	c 214	16	1.4	601	4	US-09-949-016-113977	Sequence 113977, A
c 142	17	1.4	363032	4	US-09-949-016-15754	Sequence 15754, A	c 215	16	1.4	601	4	US-09-949-016-114143	Sequence 114143, A
c 143	17	1.4	536165	4	US-09-214-808-1	Sequence 1, Appli	c 216	16	1.4	601	4	US-09-949-016-114144	Sequence 114144, A
c 144	17	1.4	1830121	4	US-09-557-884-1	Sequence 1, Appli	c 217	16	1.4	601	4	US-09-949-016-114145	Sequence 114145, A
c 145	17	1.4	1830121	4	US-09-643-990A-1	Sequence 1, Appli	c 218	16	1.4	601	4	US-09-949-016-114311	Sequence 114311, A
c 146	17	1.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli	c 219	16	1.4	601	4	US-09-949-016-114312	Sequence 114312, A
c 147	17	1.4	4403765	3	US-09-103-840A-2	Sequence 2, Appli	c 220	16	1.4	601	4	US-09-949-016-114313	Sequence 114313, A
c 148	17	1.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli	c 221	16	1.4	601	4	US-09-949-016-114479	Sequence 114479, A
c 149	17	1.4	4411529	3	US-09-103-840A-1	Sequence 1, Appli	c 222	16	1.4	601	4	US-09-949-016-114480	Sequence 114480, A
c 150	16	1.4	18	4	US-09-747-391-64	Sequence 64, Appl	c 223	16	1.4	601	4	US-09-949-016-114481	Sequence 114481, A
c 151	16	1.4	35	4	US-09-452-638-29	Sequence 29, Appl	c 224	16	1.4	601	4	US-09-949-016-114647	Sequence 114647, A
c 152	16	1.4	115	4	US-09-513-999C-29005	Sequence 29005, A	c 225	16	1.4	601	4	US-09-949-016-114648	Sequence 114648, A
c 153	16	1.4	284	4	US-08-107-532A-1304	Sequence 1304, Ap	c 226	16	1.4	601	4	US-09-949-016-114649	Sequence 114649, A
c 154	16	1.4	273	4	US-09-489-039A-2933	Sequence 2933, Ap	c 227	16	1.4	601	4	US-09-949-016-114815	Sequence 114815, A
c 155	16	1.4	331	4	US-09-489-039A-3718	Sequence 3718, Ap	c 228	16	1.4	601	4	US-09-949-016-114816	Sequence 114816, A
c 156	16	1.4	339	2	US-08-672-345C-88	Sequence 88, Appl	c 229	16	1.4	601	4	US-09-949-016-114817	Sequence 114817, A
c 157	16	1.4	339	3	US-09-214-095D-99	Sequence 99, Appl	c 230	16	1.4	601	4	US-09-949-016-114983	Sequence 114983, A
c 158	16	1.4	368	2	US-08-672-345C-86	Sequence 86, Appl	c 231	16	1.4	601	4	US-09-949-016-114984	Sequence 114984, A
c 159	16	1.4	368	2	US-08-672-345C-90	Sequence 90, Appl	c 232	16	1.4	601	4	US-09-949-016-114985	Sequence 114985, A
c 160	16	1.4	368	3	US-09-214-095D-103	Sequence 103, App	c 233	16	1.4	601	4	US-09-949-016-120449	Sequence 120449, A
c 161	16	1.4	368	3	US-09-214-095D-107	Sequence 107, App	c 234	16	1.4	601	4	US-09-949-016-120449	Sequence 120449, A
c 162	16	1.4	396	4	US-09-328-352-698	Sequence 698, App	c 235	16	1.4	601	4	US-09-949-016-122766	Sequence 122766, A
c 163	16	1.4	420	3	US-09-214-095D-111	Sequence 111, App	c 236	16	1.4	601	4	US-09-949-016-130894	Sequence 130894, A
c 164	16	1.4	450	4	US-09-489-039A-1081	Sequence 1081, Ap	c 237	16	1.4	601	4	US-09-949-016-132217	Sequence 132217, A
c 165	16	1.4	451	4	US-09-621-976-18882	Sequence 18882, A	c 238	16	1.4	601	4	US-09-949-016-138793	Sequence 138793, A
c 166	16	1.4	498	4	US-09-302-540-7531	Sequence 7531, Ap	c 239	16	1.4	601	4	US-09-949-016-142090	Sequence 142090, A
c 167	16	1.4	531	3	US-09-404-879A-96	Sequence 96, Appl	c 240	16	1.4	601	4	US-09-949-016-151285	Sequence 151285, A
c 168	16	1.4	531	4	US-09-338-933-96	Sequence 96, Appl	c 241	16	1.4	601	4	US-09-949-016-151348	Sequence 151348, A
c 169	16	1.4	531	4	US-09-215-681-96	Sequence 96, Appl	c 242	16	1.4	601	4	US-09-949-016-153542	Sequence 153542, A
c 170	16	1.4	531	4	US-09-216-003A-96	Sequence 96, Appl	c 243	16	1.4	601	4	US-09-949-016-162299	Sequence 162299, A
c 171	16	1.4	531	4	US-09-667-857-96	Sequence 96, Appl	c 244	16	1.4	601	4	US-09-949-016-162300	Sequence 162300, A
c 172	16	1.4	570	4	US-08-489-039A-3892	Sequence 3892, Ap	c 245	16	1.4	601	4	US-09-949-016-162314	Sequence 162314, A
c 173	16	1.4	573	3	US-08-906-156A-67	Sequence 67, Appl	c 246	16	1.4	601	4	US-09-949-016-162934	Sequence 162934, A

247	1.4	601	4	US-09-949-016-189935	Sequence 182935,	C 320	16	1.4	3760	4	US-09-902-540-543	Sequence 543, App
248	1.4	601	4	US-09-949-016-189373	Sequence 188973,	C 321	16	1.4	3806	4	US-09-620-312D-608	Sequence 608, App
249	1.4	601	4	US-09-949-016-189974	Sequence 189974,	C 322	16	1.4	4065	4	US-09-016-434-1105	Sequence 1105, App
250	1.4	601	4	US-09-949-016-189250	Sequence 198250,	C 323	16	1.4	4287	1	US-08-244-189-1	Sequence 1, Appli
251	1.4	601	4	US-09-949-016-200223	Sequence 200223,	C 324	16	1.4	4287	1	US-08-306-691B-53	Sequence 53, Appli
252	1.4	601	4	US-09-949-016-200224	Sequence 200224,	C 325	16	1.4	4724	4	US-09-949-016-681	Sequence 681, App
253	1.4	601	4	US-09-949-016-202719	Sequence 202719,	C 326	16	1.4	4724	4	US-09-949-016-681	Sequence 1, Appli
254	1.4	601	4	US-09-949-016-202720	Sequence 202720,	C 327	16	1.4	4739	3	US-08-685-871-1	Sequence 1, Appli
255	1.4	601	4	US-09-949-016-202720	Sequence 202720,	C 328	16	1.4	4892	4	US-09-620-312D-492	Sequence 492, App
256	1.4	601	4	US-09-949-016-205547	Sequence 205547,	C 329	16	1.4	4936	4	US-09-809-665A-29	Sequence 29, Appl
257	1.4	639	3	US-09-107-532A-1561	Sequence 1561, Ap	C 330	16	1.4	5262	4	US-09-573-080A-40	Sequence 40, Appl
258	1.4	734	3	US-08-936-165A-51	Sequence 51, Appli	C 331	16	1.4	5590	3	US-09-050-159-129	Sequence 129, App
259	1.4	801	3	US-09-459-956-7	Sequence 7, Appli	C 332	16	1.4	5703	3	US-09-280-590A-36	Sequence 36, Appl
260	1.4	825	4	US-09-902-540-3887	Sequence 3887, Ap	C 333	16	1.4	5703	4	US-09-892-398-36	Sequence 36, Appl
261	1.4	867	4	US-09-489-039A-1748	Sequence 1748, Ap	C 334	16	1.4	6008	4	US-09-949-016-4542	Sequence 4542, Ap
262	1.4	870	4	US-09-543-681A-2741	Sequence 2741, Ap	C 335	16	1.4	6152	3	US-08-973-462-1	Sequence 1, Appli
263	1.4	922	3	US-09-609-816-2	Sequence 2, Appli	C 336	16	1.4	6915	4	US-09-902-540-837	Sequence 837, App
264	1.4	924	4	US-09-902-540-4634	Sequence 4634, Ap	C 337	16	1.4	6988	4	US-09-902-540-733	Sequence 733, App
265	1.4	942	3	US-09-609-816-1	Sequence 1, Appli	C 338	16	1.4	7213	4	US-09-634-238-20	Sequence 20, Appl
266	1.4	999	4	US-09-252-991A-5206	Sequence 5206, Ap	C 339	16	1.4	7411	4	US-09-634-238-27	Sequence 27, Appl
267	1.4	1004	3	US-09-609-816-3	Sequence 3, Appli	C 340	16	1.4	7845	4	US-09-949-016-14467	Sequence 14467, A
268	1.4	1038	4	US-09-543-681A-44	Sequence 44, Appli	C 341	16	1.4	8413	4	US-09-949-016-17440	Sequence 17440, A
269	1.4	1119	4	US-09-489-039A-3603	Sequence 3603, Ap	C 342	16	1.4	8679	4	US-09-949-016-17470	Sequence 17470, A
270	1.4	1170	4	US-09-489-039A-936	Sequence 936, App	C 343	16	1.4	9509	4	US-09-949-016-15656	Sequence 15656, A
271	1.4	1173	4	US-09-107-532A-2315	Sequence 2315, Ap	C 344	16	1.4	9871	4	US-09-949-016-15470	Sequence 15470, A
272	1.4	1235	4	US-09-949-016-3914	Sequence 3914, Ap	C 345	16	1.4	10302	4	US-09-949-016-16708	Sequence 16708, A
273	1.4	1239	4	US-09-602-787A-269	Sequence 269, App	C 346	16	1.4	11349	4	US-09-949-016-16932	Sequence 16932, A
274	1.4	1254	4	US-09-248-796A-3831	Sequence 3831, Ap	C 347	16	1.4	11382	4	US-09-902-540-904	Sequence 904, App
275	1.4	1344	4	US-09-489-039A-3422	Sequence 3422, Ap	C 348	16	1.4	15368	4	US-09-949-016-13672	Sequence 13672, A
276	1.4	1365	4	US-09-252-991A-16531	Sequence 16531, A	C 349	16	1.4	16573	4	US-09-949-016-14876	Sequence 14876, A
277	1.4	1368	1	US-08-290-448A-79	Sequence 79, Appli	C 350	16	1.4	16965	4	US-09-949-016-16705	Sequence 16705, A
278	1.4	1368	1	US-08-290-448A-79	Sequence 79, Appli	C 351	16	1.4	18947	4	US-09-949-016-15106	Sequence 15106, A
279	1.4	1368	1	US-08-175-069A-79	Sequence 79, Appli	C 352	16	1.4	19736	3	US-09-740-035-3	Sequence 3, Appli
280	1.4	1368	3	US-08-461-939B-79	Sequence 79, Appli	C 353	16	1.4	20933	3	US-09-949-016-15207	Sequence 7, Appli
281	1.4	1435	4	US-09-252-991A-10987	Sequence 10987, A	C 354	16	1.4	20165	3	US-09-609-816-7	Sequence 7, Appli
282	1.4	1436	4	US-09-668-262A-11	Sequence 11, Appli	C 355	16	1.4	21143	4	US-09-902-540-1191	Sequence 1191, Ap
283	1.4	1436	4	US-09-270-767-15085	Sequence 15085, A	C 356	16	1.4	21384	4	US-09-949-016-16283	Sequence 16283, A
284	1.4	1463	4	US-09-252-991A-11262	Sequence 11262, A	C 357	16	1.4	25249	4	US-09-949-016-17444	Sequence 17444, A
285	1.4	1470	4	US-09-530-851-1	Sequence 1, Appli	C 358	16	1.4	26930	4	US-09-902-540-1228	Sequence 1228, Ap
286	1.4	1506	4	US-09-530-851-1	Sequence 1, Appli	C 359	16	1.4	27587	4	US-09-949-016-12185	Sequence 12185, A
287	1.4	1527	4	US-09-489-039A-6377	Sequence 6377, Ap	C 360	16	1.4	27589	4	US-09-949-016-17180	Sequence 17180, A
288	1.4	1539	4	US-09-107-532A-1608	Sequence 1608, Ap	C 361	16	1.4	29558	4	US-09-949-016-15607	Sequence 15607, A
289	1.4	1627	4	US-09-270-767-10633	Sequence 10633, A	C 362	16	1.4	29177	4	US-09-949-016-16284	Sequence 16284, A
290	1.4	1638	4	US-10-237-551-119	Sequence 119, App	C 363	16	1.4	29769	4	US-09-949-016-16422	Sequence 16422, A
291	1.4	1638	4	US-10-237-551-148	Sequence 148, App	C 364	16	1.4	30324	4	US-09-949-016-16037	Sequence 16037, A
292	1.4	1644	4	US-10-237-551-213	Sequence 213, App	C 365	16	1.4	30755	4	US-09-949-016-16937	Sequence 16937, A
293	1.4	1644	4	US-10-237-551-214	Sequence 214, App	C 366	16	1.4	33519	4	US-09-949-016-17165	Sequence 17165, A
294	1.4	1655	3	US-09-321-017B-531	Sequence 531, App	C 367	16	1.4	33639	4	US-09-949-016-12514	Sequence 12514, A
295	1.4	1710	4	US-09-252-991A-5271	Sequence 5271, Ap	C 368	16	1.4	33639	4	US-09-949-016-12514	Sequence 12514, A
296	1.4	1839	4	US-09-489-039A-3531	Sequence 3531, Ap	C 369	16	1.4	36203	4	US-09-949-016-15430	Sequence 15430, A
297	1.4	1909	3	US-09-100-193-6	Sequence 6, Appli	C 370	16	1.4	37059	4	US-09-949-016-15865	Sequence 15865, A
298	1.4	1952	2	US-08-392-806A-5	Sequence 5, Appli	C 371	16	1.4	38119	4	US-09-949-016-17503	Sequence 17503, A
299	1.4	1952	3	US-09-257-490-5	Sequence 5, Appli	C 372	16	1.4	38298	4	US-09-949-016-16650	Sequence 16650, A
300	1.4	1955	4	US-09-949-016-3134	Sequence 3134, Ap	C 373	16	1.4	39318	4	US-09-949-016-13798	Sequence 13798, A
301	1.4	2007	4	US-08-829-402-1	Sequence 1, Appli	C 374	16	1.4	41380	4	US-09-949-016-15128	Sequence 15128, A
302	1.4	2030	4	US-09-530-851-3	Sequence 3, Appli	C 375	16	1.4	41380	4	US-09-949-016-15471	Sequence 15471, A
303	1.4	2034	4	US-09-540-236-1780	Sequence 1780, Ap	C 376	16	1.4	41736	4	US-09-949-016-17091	Sequence 17091, A
304	1.4	2046	4	US-09-252-991A-5266	Sequence 5266, Ap	C 377	16	1.4	41737	4	US-09-949-016-12204	Sequence 12204, A
305	1.4	2121	4	US-09-614-221A-235	Sequence 235, App	C 378	16	1.4	41741	4	US-09-949-016-16983	Sequence 16983, A
306	1.4	2276	4	US-09-620-312D-612	Sequence 612, App	C 379	16	1.4	42235	4	US-09-949-016-12558	Sequence 12558, A
307	1.4	2286	4	US-09-620-312D-49	Sequence 49, Appli	C 380	16	1.4	42394	4	US-09-949-016-12752	Sequence 12752, A
308	1.4	2314	3	US-08-105-454-1	Sequence 1, Appli	C 381	16	1.4	42395	4	US-09-949-016-15115	Sequence 15115, A
309	1.4	2314	4	US-09-949-016-772	Sequence 772, App	C 382	16	1.4	43657	4	US-09-949-016-13777	Sequence 13777, A
310	1.4	2339	3	US-09-268-140-11	Sequence 11, Appli	C 383	16	1.4	44064	4	US-09-949-016-12015	Sequence 12015, A
311	1.4	2430	4	US-09-252-991A-16171	Sequence 16171, A	C 384	16	1.4	44072	4	US-09-949-016-15757	Sequence 15757, A
312	1.4	2505	3	US-09-268-140-1	Sequence 1, Appli	C 385	16	1.4	45299	4	US-09-949-016-12465	Sequence 12465, A
313	1.4	2517	3	US-09-268-140-7	Sequence 7, Appli	C 386	16	1.4	45299	4	US-09-949-016-13945	Sequence 13945, A
314	1.4	2592	4	US-09-489-039A-3913	Sequence 3913, Ap	C 387	16	1.4	47799	4	US-09-949-016-13363	Sequence 13363, A
315	1.4	2771	4	US-09-807-757C-7	Sequence 7, Appli	C 388	16	1.4	47964	4	US-09-949-016-13363	Sequence 13363, A
316	1.4	2870	1	US-08-468-036-28	Sequence 28, Appli	C 389	16	1.4	49164	4	US-09-949-016-12985	Sequence 12985, A
317	1.4	3213	2	US-08-376-843-28	Sequence 28, Appli	C 390	16	1.4	49164	4	US-09-949-016-12986	Sequence 12986, A
318	1.4	3270	4	US-09-949-016-5184	Sequence 5184, Ap	C 391	16	1.4	49164	4	US-09-949-016-13262	Sequence 13262, A
319	1.4	3758	4	US-09-902-540-6742	Sequence 6742, Ap	C 392	16	1.4	49164	4	US-09-949-016-13263	Sequence 13263, A

539	15	1.3	284	3	US-08-905-223-219	Sequence 219, App	c 612	15	1.3	523	4	US-09-270-767-1684	Sequence 1684, Ap
540	15	1.3	288	4	US-09-313-294A-6282	Sequence 6282, Ap	c 613	15	1.3	523	4	US-09-270-767-16966	Sequence 16966, A
541	15	1.3	288	4	US-09-489-039A-3005	Sequence 3005, Ap	c 614	15	1.3	540	3	US-09-221-017B-274	Sequence 274, App
542	15	1.3	293	4	US-09-621-976-15168	Sequence 15168, A	615	15	1.3	547	4	US-09-389-681-450	Sequence 450, App
543	15	1.3	303	3	US-09-247-527-1	Sequence 1, Appl	616	15	1.3	547	4	US-09-620-405B-450	Sequence 450, App
544	15	1.3	312	4	US-09-489-039A-6509	Sequence 6509, App	617	15	1.3	547	4	US-09-433-826B-450	Sequence 450, App
545	15	1.3	326	4	US-09-270-767-25371	Sequence 25371, A	618	15	1.3	547	4	US-09-604-287A-450	Sequence 450, App
546	15	1.3	328	4	US-09-513-999C-12733	Sequence 12733, A	619	15	1.3	547	4	US-09-834-759-450	Sequence 450, App
547	15	1.3	335	4	US-09-513-999C-34144	Sequence 34144, A	620	15	1.3	547	4	US-09-590-751A-450	Sequence 450, App
548	15	1.3	340	4	US-09-270-767-5420	Sequence 5420, Ap	621	15	1.3	547	4	US-09-551-621-450	Sequence 450, App
549	15	1.3	340	4	US-09-270-767-20702	Sequence 20702, Ap	622	15	1.3	548	4	US-09-621-976-2414	Sequence 2414, Ap
550	15	1.3	344	4	US-09-513-999C-35482	Sequence 35482, A	c 623	15	1.3	550	4	US-09-949-016-184384	Sequence 184384,
551	15	1.3	345	4	US-09-540-236-558	Sequence 558, App	c 624	15	1.3	552	4	US-08-356-171B-1109	Sequence 1109, Ap
552	15	1.3	350	4	US-09-513-999C-2526	Sequence 2526, App	c 625	15	1.3	552	4	US-09-621-976-18669	Sequence 18669, A
553	15	1.3	363	4	US-09-270-767-390	Sequence 390, App	c 626	15	1.3	552	4	US-08-781-986A-1109	Sequence 1109, Ap
554	15	1.3	363	4	US-09-270-767-15672	Sequence 15672, A	c 627	15	1.3	576	3	US-09-134-001C-1610	Sequence 1610, Ap
555	15	1.3	369	3	US-09-643-597-148	Sequence 148, App	628	15	1.3	576	3	US-09-134-000C-3351	Sequence 3351, Ap
556	15	1.3	369	4	US-09-480-884A-148	Sequence 148, App	c 629	15	1.3	582	4	US-09-621-976-3834	Sequence 3834, Ap
557	15	1.3	369	4	US-09-542-615A-148	Sequence 148, App	c 630	15	1.3	592	4	US-09-248-796A-2750	Sequence 2750, Ap
558	15	1.3	369	4	US-09-606-421B-148	Sequence 148, App	c 631	15	1.3	600	4	US-09-489-039A-1097	Sequence 1097, Ap
559	15	1.3	369	4	US-09-221-107-148	Sequence 148, App	c 632	15	1.3	601	4	US-09-949-016-18080	Sequence 18080, A
560	15	1.3	369	4	US-09-466-396A-148	Sequence 148, App	c 633	15	1.3	601	4	US-09-949-016-21559	Sequence 21559, A
561	15	1.3	369	4	US-09-476-496A-148	Sequence 148, App	c 634	15	1.3	601	4	US-09-949-016-23287	Sequence 23287, A
562	15	1.3	369	4	US-09-630-940B-148	Sequence 148, App	c 635	15	1.3	601	4	US-09-949-016-23442	Sequence 23442, A
563	15	1.3	369	4	US-09-285-479-148	Sequence 148, App	c 636	15	1.3	601	4	US-09-949-016-28609	Sequence 28609, A
564	15	1.3	372	4	US-09-248-796A-9978	Sequence 9978, Ap	c 637	15	1.3	601	4	US-09-949-016-28610	Sequence 28610, A
565	15	1.3	373	4	US-09-621-976-16725	Sequence 16725, A	c 638	15	1.3	601	4	US-09-949-016-30598	Sequence 30598, A
566	15	1.3	381	4	US-09-513-999C-10526	Sequence 10526, A	c 639	15	1.3	601	4	US-09-949-016-30599	Sequence 30599, A
567	15	1.3	392	4	US-09-640-211A-1553	Sequence 1553, Ap	c 640	15	1.3	601	4	US-09-949-016-32241	Sequence 32241, A
568	15	1.3	396	4	US-09-513-999C-2992	Sequence 2992, Ap	641	15	1.3	601	4	US-09-949-016-32677	Sequence 32677, A
569	15	1.3	397	4	US-09-270-767-31323	Sequence 31323, A	642	15	1.3	601	4	US-09-949-016-32678	Sequence 32678, A
570	15	1.3	397	4	US-09-513-999C-21697	Sequence 21697, A	643	15	1.3	601	4	US-09-949-016-33137	Sequence 33137, A
571	15	1.3	407	3	US-09-615-192A-204	Sequence 204, App	644	15	1.3	601	4	US-09-949-016-33138	Sequence 33138, A
572	15	1.3	408	4	US-09-513-999C-34511	Sequence 34511, A	645	15	1.3	601	4	US-09-949-016-33139	Sequence 33139, A
573	15	1.3	417	4	US-09-328-352-2570	Sequence 2570, Ap	c 646	15	1.3	601	4	US-09-949-016-34696	Sequence 34696, A
574	15	1.3	419	4	US-09-270-767-5642	Sequence 5642, Ap	c 647	15	1.3	601	4	US-09-949-016-34697	Sequence 34697, A
575	15	1.3	419	4	US-09-270-767-20924	Sequence 20924, Ap	c 648	15	1.3	601	4	US-09-949-016-36389	Sequence 36389, A
576	15	1.3	424	4	US-09-270-767-10670	Sequence 10670, A	649	15	1.3	601	4	US-09-949-016-36390	Sequence 36390, A
577	15	1.3	431	1	US-08-474-140-15	Sequence 15, Appl	650	15	1.3	601	4	US-09-949-016-36391	Sequence 36391, A
578	15	1.3	431	1	US-08-477-630-15	Sequence 15, Appl	651	15	1.3	601	4	US-09-949-016-36408	Sequence 36408, A
579	15	1.3	431	1	US-08-472-293-15	Sequence 15, Appl	652	15	1.3	601	4	US-09-949-016-36409	Sequence 36409, A
580	15	1.3	431	1	US-08-474-545-15	Sequence 15, Appl	653	15	1.3	601	4	US-09-949-016-36410	Sequence 36410, A
581	15	1.3	431	1	US-08-478-341-15	Sequence 15, Appl	c 654	15	1.3	601	4	US-09-949-016-39411	Sequence 39411, A
582	15	1.3	431	3	US-08-996-733-15	Sequence 15, Appl	c 655	15	1.3	601	4	US-09-949-016-40086	Sequence 40086, A
583	15	1.3	431	4	US-09-894-882-73	Sequence 73, Appl	656	15	1.3	601	4	US-09-949-016-40710	Sequence 40710, A
584	15	1.3	431	4	US-09-513-999C-36573	Sequence 36573, A	c 657	15	1.3	601	4	US-09-949-016-44689	Sequence 44689, A
585	15	1.3	435	4	US-10-101-464A-240	Sequence 240, App	c 658	15	1.3	601	4	US-09-949-016-44690	Sequence 44690, A
586	15	1.3	440	4	US-09-621-976-18472	Sequence 18472, A	c 659	15	1.3	601	4	US-09-949-016-47043	Sequence 47043, A
587	15	1.3	454	4	US-09-513-999C-1121	Sequence 1121, Ap	c 660	15	1.3	601	4	US-09-949-016-47044	Sequence 47044, A
588	15	1.3	455	4	US-09-702-705-58	Sequence 58, Appl	c 661	15	1.3	601	4	US-09-949-016-53565	Sequence 53565, A
589	15	1.3	455	4	US-09-736-457-58	Sequence 58, Appl	c 662	15	1.3	601	4	US-09-949-016-53689	Sequence 53689, A
590	15	1.3	455	4	US-09-614-124B-58	Sequence 58, Appl	663	15	1.3	601	4	US-09-949-016-55706	Sequence 55706, A
591	15	1.3	455	4	US-09-671-325-58	Sequence 58, Appl	664	15	1.3	601	4	US-09-949-016-55706	Sequence 55706, A
592	15	1.3	455	4	US-09-589-184-58	Sequence 58, Appl	665	15	1.3	601	4	US-09-949-016-58012	Sequence 58012, A
593	15	1.3	455	4	US-09-658-824-58	Sequence 58, Appl	c 666	15	1.3	601	4	US-09-949-016-59795	Sequence 59795, A
594	15	1.3	455	4	US-09-513-999C-10525	Sequence 10525, A	667	15	1.3	601	4	US-09-949-016-60208	Sequence 60208, A
595	15	1.3	456	3	US-08-704-966-8	Sequence 8, Appl	668	15	1.3	601	4	US-09-949-016-60814	Sequence 60814, A
596	15	1.3	456	3	US-08-705-438-8	Sequence 8, Appl	669	15	1.3	601	4	US-09-949-016-60815	Sequence 60815, A
597	15	1.3	456	4	US-09-328-352-2410	Sequence 2410, Ap	670	15	1.3	601	4	US-09-949-016-60816	Sequence 60816, A
598	15	1.3	483	3	US-09-134-001C-2017	Sequence 2017, Ap	671	15	1.3	601	4	US-09-949-016-64999	Sequence 64999, A
599	15	1.3	484	4	US-09-621-976-12877	Sequence 12877, A	c 672	15	1.3	601	4	US-09-949-016-66663	Sequence 66663, A
600	15	1.3	485	3	US-09-621-976-16723	Sequence 16723, A	c 673	15	1.3	601	4	US-09-949-016-68720	Sequence 68720, A
601	15	1.3	486	3	US-09-134-001C-1428	Sequence 1428, Ap	674	15	1.3	601	4	US-09-949-016-69358	Sequence 69358, A
602	15	1.3	486	4	US-09-252-991A-1006	Sequence 1006, Ap	675	15	1.3	601	4	US-09-949-016-69359	Sequence 69359, A
603	15	1.3	500	4	US-09-799-451-335	Sequence 335, App	676	15	1.3	601	4	US-09-949-016-69360	Sequence 69360, A
604	15	1.3	505	4	US-09-621-976-3341	Sequence 3341, Ap	677	15	1.3	601	4	US-09-949-016-69361	Sequence 69361, A
605	15	1.3	507	3	US-09-328-111-826	Sequence 111, App	c 678	15	1.3	601	4	US-09-949-016-71558	Sequence 71558, A
606	15	1.3	508	4	US-09-621-976-1387	Sequence 1387, App	679	15	1.3	601	4	US-09-949-016-72639	Sequence 72639, A
607	15	1.3	519	4	US-09-621-976-1818	Sequence 1818, Ap	680	15	1.3	601	4	US-09-949-016-72640	Sequence 72640, A
608	15	1.3	520	4	US-09-023-655-402	Sequence 402, App	c 681	15	1.3	601	4	US-09-949-016-73384	Sequence 73384, A
609	15	1.3	522	4	US-09-949-016-184385	Sequence 184385,	c 682	15	1.3	601	4	US-09-949-016-73385	Sequence 73385, A
610	15	1.3	523	3	US-09-196-387-12	Sequence 12, Appl	c 683	15	1.3	601	4	US-09-949-016-73626	Sequence 73626, A
611	15	1.3	523	4	US-09-841-835-12	Sequence 12, Appl	684	15	1.3	601	4	US-09-949-016-77042	Sequence 77042, A

c 685	15	1.3	601	4	US-09-949-016-77392	Sequence 77392, A	758	15	1.3	601	4	US-09-949-016-150818	Sequence 150818,
c 686	15	1.3	601	4	US-09-949-016-77393	Sequence 77393, A	759	15	1.3	601	4	US-09-949-016-151315	Sequence 151315,
c 687	15	1.3	601	4	US-09-949-016-78470	Sequence 78470, A	760	15	1.3	601	4	US-09-949-016-151316	Sequence 151316,
c 688	15	1.3	601	4	US-09-949-016-80114	Sequence 80114, A	c 761	15	1.3	601	4	US-09-949-016-151324	Sequence 151324,
c 689	15	1.3	601	4	US-09-949-016-80115	Sequence 80115, A	c 762	15	1.3	601	4	US-09-949-016-151325	Sequence 151325,
c 690	15	1.3	601	4	US-09-949-016-80721	Sequence 80721, A	c 763	15	1.3	601	4	US-09-949-016-151378	Sequence 151378,
c 691	15	1.3	601	4	US-09-949-016-80722	Sequence 80722, A	c 764	15	1.3	601	4	US-09-949-016-151379	Sequence 151379,
c 692	15	1.3	601	4	US-09-949-016-85479	Sequence 85479, A	c 765	15	1.3	601	4	US-09-949-016-151387	Sequence 151387,
c 693	15	1.3	601	4	US-09-949-016-85821	Sequence 85821, A	c 766	15	1.3	601	4	US-09-949-016-151388	Sequence 151388,
c 694	15	1.3	601	4	US-09-949-016-86082	Sequence 86082, A	c 767	15	1.3	601	4	US-09-949-016-152511	Sequence 152511,
c 695	15	1.3	601	4	US-09-949-016-86083	Sequence 86083, A	c 768	15	1.3	601	4	US-09-949-016-153206	Sequence 153206,
c 696	15	1.3	601	4	US-09-949-016-86546	Sequence 86546, A	c 769	15	1.3	601	4	US-09-949-016-153951	Sequence 153951,
c 697	15	1.3	601	4	US-09-949-016-87537	Sequence 87537, A	c 770	15	1.3	601	4	US-09-949-016-153952	Sequence 153952,
c 698	15	1.3	601	4	US-09-949-016-88857	Sequence 88857, A	c 771	15	1.3	601	4	US-09-949-016-153952	Sequence 153952,
c 699	15	1.3	601	4	US-09-949-016-88858	Sequence 88858, A	c 772	15	1.3	601	4	US-09-949-016-15421	Sequence 15421,
c 700	15	1.3	601	4	US-09-949-016-89325	Sequence 89325, A	c 773	15	1.3	601	4	US-09-949-016-154213	Sequence 154213,
c 701	15	1.3	601	4	US-09-949-016-91959	Sequence 91959, A	c 774	15	1.3	601	4	US-09-949-016-159214	Sequence 159214,
c 702	15	1.3	601	4	US-09-949-016-93974	Sequence 93974, A	c 775	15	1.3	601	4	US-09-949-016-162528	Sequence 162528,
c 703	15	1.3	601	4	US-09-949-016-93975	Sequence 93975, A	c 776	15	1.3	601	4	US-09-949-016-162528	Sequence 162528,
c 704	15	1.3	601	4	US-09-949-016-95981	Sequence 95981, A	c 777	15	1.3	601	4	US-09-949-016-165692	Sequence 165692,
c 705	15	1.3	601	4	US-09-949-016-103408	Sequence 103408, A	c 778	15	1.3	601	4	US-09-949-016-166060	Sequence 166060,
c 706	15	1.3	601	4	US-09-949-016-105096	Sequence 105096, A	c 779	15	1.3	601	4	US-09-949-016-168859	Sequence 168859,
c 707	15	1.3	601	4	US-09-949-016-105097	Sequence 105097, A	c 780	15	1.3	601	4	US-09-949-016-168860	Sequence 168860,
c 708	15	1.3	601	4	US-09-949-016-105826	Sequence 105826, A	c 781	15	1.3	601	4	US-09-949-016-169617	Sequence 169617,
c 709	15	1.3	601	4	US-09-949-016-105827	Sequence 105827, A	c 782	15	1.3	601	4	US-09-949-016-170109	Sequence 170109,
c 710	15	1.3	601	4	US-09-949-016-105828	Sequence 105828, A	c 783	15	1.3	601	4	US-09-949-016-170540	Sequence 170540,
c 711	15	1.3	601	4	US-09-949-016-105840	Sequence 105840, A	c 784	15	1.3	601	4	US-09-949-016-170541	Sequence 170541,
c 712	15	1.3	601	4	US-09-949-016-105841	Sequence 105841, A	c 785	15	1.3	601	4	US-09-949-016-173592	Sequence 173592,
c 713	15	1.3	601	4	US-09-949-016-105842	Sequence 105842, A	c 786	15	1.3	601	4	US-09-949-016-173907	Sequence 173907,
c 714	15	1.3	601	4	US-09-949-016-105854	Sequence 105854, A	c 787	15	1.3	601	4	US-09-949-016-173907	Sequence 173907,
c 715	15	1.3	601	4	US-09-949-016-105855	Sequence 105855, A	c 788	15	1.3	601	4	US-09-949-016-179609	Sequence 179609,
c 716	15	1.3	601	4	US-09-949-016-105856	Sequence 105856, A	c 789	15	1.3	601	4	US-09-949-016-181525	Sequence 181525,
c 717	15	1.3	601	4	US-09-949-016-112421	Sequence 112421, A	c 790	15	1.3	601	4	US-09-949-016-181526	Sequence 181526,
c 718	15	1.3	601	4	US-09-949-016-112422	Sequence 112422, A	c 791	15	1.3	601	4	US-09-949-016-181527	Sequence 181527,
c 719	15	1.3	601	4	US-09-949-016-116856	Sequence 116856, A	c 792	15	1.3	601	4	US-09-949-016-185884	Sequence 185884,
c 720	15	1.3	601	4	US-09-949-016-116974	Sequence 116974, A	c 793	15	1.3	601	4	US-09-949-016-188160	Sequence 188160,
c 721	15	1.3	601	4	US-09-949-016-118159	Sequence 118159, A	c 794	15	1.3	601	4	US-09-949-016-189491	Sequence 189491,
c 722	15	1.3	601	4	US-09-949-016-121013	Sequence 121013, A	c 795	15	1.3	601	4	US-09-949-016-190285	Sequence 190285,
c 723	15	1.3	601	4	US-09-949-016-121798	Sequence 121798, A	c 796	15	1.3	601	4	US-09-949-016-190286	Sequence 190286,
c 724	15	1.3	601	4	US-09-949-016-122164	Sequence 122164, A	c 797	15	1.3	601	4	US-09-949-016-193518	Sequence 193518,
c 725	15	1.3	601	4	US-09-949-016-122216	Sequence 122216, A	c 798	15	1.3	601	4	US-09-949-016-197934	Sequence 197934,
c 726	15	1.3	601	4	US-09-949-016-123911	Sequence 123911, A	c 799	15	1.3	601	4	US-09-949-016-197937	Sequence 197937,
c 727	15	1.3	601	4	US-09-949-016-125303	Sequence 125303, A	c 800	15	1.3	601	4	US-09-949-016-197998	Sequence 197998,
c 728	15	1.3	601	4	US-09-949-016-125304	Sequence 125304, A	c 801	15	1.3	601	4	US-09-949-016-198033	Sequence 198033,
c 729	15	1.3	601	4	US-09-949-016-126661	Sequence 126661, A	c 802	15	1.3	601	4	US-09-949-016-198880	Sequence 198880,
c 730	15	1.3	601	4	US-09-949-016-126662	Sequence 126662, A	c 803	15	1.3	601	4	US-09-949-016-198881	Sequence 198881,
c 731	15	1.3	601	4	US-09-949-016-127257	Sequence 127257, A	c 804	15	1.3	601	4	US-09-949-016-199220	Sequence 199220,
c 732	15	1.3	601	4	US-09-949-016-128888	Sequence 128888, A	c 805	15	1.3	601	4	US-09-949-016-199221	Sequence 199221,
c 733	15	1.3	601	4	US-09-949-016-131707	Sequence 131707, A	c 806	15	1.3	601	4	US-09-949-016-199222	Sequence 199222,
c 734	15	1.3	601	4	US-09-949-016-131760	Sequence 131760, A	c 807	15	1.3	601	4	US-09-949-016-199289	Sequence 199289,
c 735	15	1.3	601	4	US-09-949-016-132605	Sequence 132605, A	c 808	15	1.3	601	4	US-09-949-016-200100	Sequence 200100,
c 736	15	1.3	601	4	US-09-949-016-132606	Sequence 132606, A	c 809	15	1.3	601	4	US-09-949-016-200101	Sequence 200101,
c 737	15	1.3	601	4	US-09-949-016-133138	Sequence 133138, A	c 810	15	1.3	601	4	US-09-949-016-200102	Sequence 200102,
c 738	15	1.3	601	4	US-09-949-016-133221	Sequence 133221, A	c 811	15	1.3	601	4	US-09-949-016-202976	Sequence 202976,
c 739	15	1.3	601	4	US-09-949-016-135106	Sequence 135106, A	c 812	15	1.3	601	4	US-09-949-016-203881	Sequence 203881,
c 740	15	1.3	601	4	US-09-949-016-135780	Sequence 135780, A	c 813	15	1.3	601	4	US-09-949-016-203882	Sequence 203882,
c 741	15	1.3	601	4	US-09-949-016-135897	Sequence 135897, A	c 814	15	1.3	601	4	US-09-949-016-203883	Sequence 203883,
c 742	15	1.3	601	4	US-09-949-016-136014	Sequence 136014, A	c 815	15	1.3	601	4	US-09-949-016-205775	Sequence 205775,
c 743	15	1.3	601	4	US-09-949-016-140152	Sequence 140152, A	c 816	15	1.3	601	4	US-09-949-016-206026	Sequence 206026,
c 744	15	1.3	601	4	US-09-949-016-140153	Sequence 140153, A	c 817	15	1.3	601	4	US-09-949-001-119	Sequence 119, App
c 745	15	1.3	601	4	US-09-949-016-141179	Sequence 141179, A	c 818	15	1.3	601	4	US-09-949-001-140	Sequence 140, App
c 746	15	1.3	601	4	US-09-949-016-143935	Sequence 143935, A	c 819	15	1.3	601	4	US-09-949-001-575	Sequence 575, App
c 747	15	1.3	601	4	US-09-949-016-143977	Sequence 143977, A	c 820	15	1.3	601	4	US-09-949-001-606	Sequence 606, App
c 748	15	1.3	601	4	US-09-949-016-144019	Sequence 144019, A	c 821	15	1.3	609	4	US-09-252-991A-12176	Sequence 12176, A
c 749	15	1.3	601	4	US-09-949-016-145293	Sequence 145293, A	c 822	15	1.3	609	4	US-09-614-221A-314	Sequence 314, App
c 750	15	1.3	601	4	US-09-949-016-146877	Sequence 146877, A	c 823	15	1.3	610	4	US-09-270-767-15032	Sequence 15032, A
c 751	15	1.3	601	4	US-09-949-016-147917	Sequence 147917, A	c 824	15	1.3	614	4	US-09-621-976-3527	Sequence 3527, Ap
c 752	15	1.3	601	4	US-09-949-016-147918	Sequence 147918, A	c 825	15	1.3	621	4	US-09-134-000C-1726	Sequence 1726, Ap
c 753	15	1.3	601	4	US-09-949-016-147919	Sequence 147919, A	c 826	15	1.3	630	4	US-09-252-991A-14077	Sequence 14077, A
c 754	15	1.3	601	4	US-09-949-016-148058	Sequence 148058, A	c 827	15	1.3	633	4	US-09-902-540-2138	Sequence 2138, Ap
c 755	15	1.3	601	4	US-09-949-016-148059	Sequence 148059, A	c 828	15	1.3	638	4	US-09-902-540-1339	Sequence 1339, Ap
c 756	15	1.3	601	4	US-09-949-016-148164	Sequence 148164, A	c 829	15	1.3	645	4	US-09-252-991A-308	Sequence 308, App
c 757	15	1.3	601	4	US-09-949-016-148165	Sequence 148165, A	c 830	15	1.3	645	4	US-09-252-991A-14347	Sequence 14347, A

831	15	1.3	651	3	US-08-715-628B-6	Sequence 6, Appli	904	15	1.3	972	4	US-09-252-991A-5171	Sequence 5171, Ap
832	15	1.3	667	4	US-09-023-655-672	Sequence 672, App	c 905	15	1.3	975	3	US-09-134-001C-2688	Sequence 2688, Ap
c 833	15	1.3	669	4	US-09-543-681A-2579	Sequence 2579, Ap	c 906	15	1.3	991	4	US-09-620-312D-787	Sequence 787, App
834	15	1.3	678	4	US-09-949-016-861	Sequence 861, App	907	15	1.3	1014	3	US-09-252-991A-356	Sequence 356, App
835	15	1.3	678	4	US-09-949-016-2706	Sequence 2706, Ap	c 908	15	1.3	1020	3	US-09-537-522-2	Sequence 2, Appli
836	15	1.3	687	4	US-09-710-279-1165	Sequence 1165, Ap	c 909	15	1.3	1020	4	US-10-067-291-2	Sequence 2, Appli
837	15	1.3	688	6	5498694-3	Patent No. 5498694	c 910	15	1.3	1038	4	US-09-252-991A-14155	Sequence 14155, A
838	15	1.3	688	6	5498694-3	Patent No. 5498694	911	15	1.3	1059	4	US-09-949-016-5611	Sequence 5611, Ap
c 839	15	1.3	689	4	US-09-586-106D-98	Sequence 98, Appl	c 912	15	1.3	1077	4	US-09-543-681A-401	Sequence 401, App
840	15	1.3	693	4	US-09-270-767-10023	Sequence 10023, A	c 913	15	1.3	1095	3	US-08-875-811-52	Sequence 52, Appl
841	15	1.3	702	3	US-08-715-628B-4	Sequence 4, Appli	c 914	15	1.3	1095	4	US-09-743-475-2	Sequence 2, Appli
c 842	15	1.3	703	3	US-09-280-116-175	Sequence 175, App	915	15	1.3	1095	4	US-09-270-767-26255	Sequence 26255, A
843	15	1.3	708	3	US-08-715-628B-5	Sequence 5, Appli	c 916	15	1.3	1098	3	US-08-875-811-54	Sequence 54, Appl
844	15	1.3	709	3	US-09-328-111-286	Sequence 286, App	917	15	1.3	1111	4	US-09-621-976-1477	Sequence 1477, Ap
845	15	1.3	711	4	US-09-252-991A-2395	Sequence 2395, App	c 918	15	1.3	1122	4	US-09-489-039A-3211	Sequence 3211, Ap
c 846	15	1.3	717	4	US-09-489-039A-4514	Sequence 4514, Ap	919	15	1.3	1134	4	US-09-949-016-3997	Sequence 3997, Ap
c 847	15	1.3	718	3	US-09-142-078-45	Sequence 45, Appl	920	15	1.3	1137	4	US-09-107-532A-2674	Sequence 2674, Ap
c 848	15	1.3	718	3	US-09-357-141-45	Sequence 45, Appl	921	15	1.3	1149	4	US-09-252-991A-14143	Sequence 14143, A
c 849	15	1.3	718	3	US-09-533-889-45	Sequence 45, Appl	c 922	15	1.3	1159	4	US-09-270-767-9812	Sequence 9812, A
c 850	15	1.3	718	4	US-09-142-080-45	Sequence 45, Appl	c 923	15	1.3	1159	4	US-09-270-767-25094	Sequence 25094, A
851	15	1.3	729	4	US-09-248-796A-4057	Sequence 4057, Ap	c 924	15	1.3	1170	4	US-09-543-681A-289	Sequence 289, App
852	15	1.3	744	4	US-09-248-796A-5707	Sequence 5707, Ap	c 925	15	1.3	1179	3	US-08-867-611-13	Sequence 13, Appl
c 853	15	1.3	750	3	US-07-715-628B-2	Sequence 2, Appli	c 926	15	1.3	1179	4	US-09-690-359-13	Sequence 13, Appl
854	15	1.3	750	4	US-09-252-991A-13978	Sequence 13978, A	c 927	15	1.3	1179	5	PCT-US92-06965A-18	Sequence 18, Appl
c 855	15	1.3	751	3	US-08-896-164-58	Sequence 58, Appl	c 928	15	1.3	1182	4	US-09-252-991A-12239	Sequence 12239, A
c 856	15	1.3	756	4	US-09-134-000C-1434	Sequence 1434, Ap	c 929	15	1.3	1187	5	PCT-US95-13536-2	Sequence 2, Appli
c 857	15	1.3	759	4	US-09-583-110-2456	Sequence 2456, Ap	c 930	15	1.3	1188	5	PCT-US95-13536-1	Sequence 1, Appli
c 858	15	1.3	765	4	US-09-270-767-12848	Sequence 12848, A	c 931	15	1.3	1197	4	US-09-710-279-43	Sequence 43, Appl
c 859	15	1.3	771	4	US-09-252-991A-5075	Sequence 5075, Ap	c 932	15	1.3	1197	4	US-09-710-279-1497	Sequence 1497, Ap
c 860	15	1.3	780	4	US-09-252-991A-12108	Sequence 12108, A	c 933	15	1.3	1207	3	US-09-264-419C-1	Sequence 1, Appli
c 861	15	1.3	785	3	US-08-836-236-3	Sequence 3, Appli	c 934	15	1.3	1209	4	US-09-248-796A-5395	Sequence 5395, Ap
c 862	15	1.3	785	4	US-09-535-679-3	Sequence 3, Appli	c 935	15	1.3	1227	4	US-09-248-796A-12443	Sequence 12443, A
c 863	15	1.3	792	4	US-09-902-540-6420	Sequence 6420, Ap	c 936	15	1.3	1243	4	US-09-806-536A-19	Sequence 19, Appl
c 864	15	1.3	798	4	US-09-252-991A-6036	Sequence 6036, Ap	c 937	15	1.3	1248	4	US-09-489-039A-6144	Sequence 6144, Ap
c 865	15	1.3	816	4	US-09-270-767-8578	Sequence 8578, Ap	c 938	15	1.3	1254	3	US-09-134-001C-973	Sequence 973, App
c 866	15	1.3	816	4	US-09-270-767-23860	Sequence 23860, A	c 939	15	1.3	1269	4	US-09-252-991A-8698	Sequence 8698, App
c 867	15	1.3	821	3	US-09-361-707-109	Sequence 109, App	940	15	1.3	1274	4	US-09-148-545-72	Sequence 72, Appl
c 868	15	1.3	822	4	US-09-252-991A-3058	Sequence 3058, Ap	c 941	15	1.3	1274	4	US-09-949-016-5715	Sequence 5715, Ap
c 869	15	1.3	822	4	US-09-107-433-1681	Sequence 1681, Ap	942	15	1.3	1296	4	US-09-148-545-126	Sequence 126, App
c 870	15	1.3	828	4	US-09-583-110-345	Sequence 345, App	943	15	1.3	1310	3	US-09-047-288-1	Sequence 1, Appli
c 871	15	1.3	831	4	US-09-540-236-444	Sequence 444, App	944	15	1.3	1310	3	US-08-802-191-1	Sequence 1, Appli
c 872	15	1.3	831	4	US-09-107-433-466	Sequence 466, App	945	15	1.3	1320	4	US-09-134-000C-2005	Sequence 2005, Ap
c 873	15	1.3	840	3	US-09-264-419C-3	Sequence 3, Appli	946	15	1.3	1322	4	US-09-270-767-13847	Sequence 13847, A
c 874	15	1.3	843	4	US-09-205-258-14	Sequence 14, Appl	c 947	15	1.3	1323	4	US-09-489-039A-5432	Sequence 5432, Ap
c 875	15	1.3	845	4	US-09-949-016-3546	Sequence 3546, Ap	c 948	15	1.3	1332	4	US-09-328-352-2817	Sequence 2817, Ap
c 876	15	1.3	846	3	US-09-455-960-5	Sequence 5, Appli	c 949	15	1.3	1335	4	US-09-328-352-2618	Sequence 2618, Ap
c 877	15	1.3	846	4	US-10-051-325-5	Sequence 5, Appli	c 950	15	1.3	1335	4	US-09-640-211A-232	Sequence 232, App
c 878	15	1.3	849	4	US-09-270-767-2089	Sequence 2089, Ap	c 951	15	1.3	1352	4	US-09-543-681A-1187	Sequence 1187, Ap
c 879	15	1.3	849	4	US-09-270-767-17371	Sequence 17371, A	c 952	15	1.3	1368	4	US-09-540-236-1041	Sequence 1041, Ap
c 880	15	1.3	851	3	US-09-126-646-3	Sequence 3, Appli	c 953	15	1.3	1371	4	US-09-134-000C-1998	Sequence 1998, Ap
c 881	15	1.3	851	3	US-09-421-491-3	Sequence 3, Appli	954	15	1.3	1377	4	US-09-252-991A-1120	Sequence 1120, Ap
c 882	15	1.3	861	3	US-09-134-001C-1266	Sequence 1266, Ap	c 955	15	1.3	1398	4	US-09-489-039A-986	Sequence 986, App
c 883	15	1.3	864	4	US-09-976-594-1054	Sequence 1054, Ap	c 956	15	1.3	1398	4	US-09-134-000C-3020	Sequence 3020, Ap
c 884	15	1.3	864	4	US-09-919-039-385	Sequence 385, App	c 957	15	1.3	1410	4	US-09-252-991A-6871	Sequence 6871, Ap
c 885	15	1.3	882	1	US-08-392-419-3	Sequence 3, Appli	958	15	1.3	1432	4	US-09-620-312D-976	Sequence 976, App
c 886	15	1.3	885	4	US-09-540-236-471	Sequence 471, App	959	15	1.3	1451	4	US-09-949-016-2143	Sequence 2143, Ap
c 887	15	1.3	888	4	US-09-252-991A-13993	Sequence 13993, A	c 960	15	1.3	1453	2	US-08-169-948B-9	Sequence 9, Appli
c 888	15	1.3	888	4	US-09-902-540-3358	Sequence 3358, Ap	c 961	15	1.3	1453	2	US-08-448-873-9	Sequence 9, Appli
c 889	15	1.3	897	4	US-09-489-039A-4580	Sequence 4580, Ap	c 962	15	1.3	1453	3	US-08-382-452D-9	Sequence 9, Appli
c 890	15	1.3	907	4	US-09-023-655-682	Sequence 682, App	c 963	15	1.3	1453	3	US-09-916-494A-9	Sequence 9, Appli
c 891	15	1.3	918	4	US-09-710-279-1263	Sequence 1263, Ap	c 964	15	1.3	1462	3	US-08-806-743-3	Sequence 3, Appli
c 892	15	1.3	921	4	US-09-248-796A-3847	Sequence 3847, Ap	c 965	15	1.3	1470	4	US-09-949-016-186	Sequence 186, App
c 893	15	1.3	933	4	US-09-902-540-8434	Sequence 8434, Ap	c 966	15	1.3	1482	4	US-09-252-991A-6850	Sequence 6850, Ap
c 894	15	1.3	934	3	US-09-221-017B-796	Sequence 796, App	c 967	15	1.3	1485	3	US-09-134-000C-2345	Sequence 2345, Ap
c 895	15	1.3	934	4	US-09-640-211A-2013	Sequence 2013, Ap	c 968	15	1.3	1491	3	US-09-643-597-171	Sequence 171, App
c 896	15	1.3	946	3	US-08-924-747-11	Sequence 11, Appl	c 969	15	1.3	1491	4	US-09-480-884A-171	Sequence 171, App
c 897	15	1.3	946	3	US-09-247-373B-11	Sequence 11, Appl	c 970	15	1.3	1491	4	US-09-542-615A-171	Sequence 171, App
c 898	15	1.3	946	3	US-09-236-715-11	Sequence 11, Appl	c 971	15	1.3	1491	4	US-09-606-421B-171	Sequence 171, App
c 899	15	1.3	948	3	US-09-134-001C-1543	Sequence 1543, Ap	c 972	15	1.3	1491	4	US-09-466-396A-171	Sequence 171, App
c 900	15	1.3	951	4	US-09-248-796A-11699	Sequence 11699, A	c 973	15	1.3	1491	4	US-09-476-496A-171	Sequence 171, App
c 901	15	1.3	965	4	US-09-583-110-459	Sequence 459, App	c 974	15	1.3	1491	4	US-09-630-940B-171	Sequence 171, App
c 902	15	1.3	969	4	US-09-107-433-452	Sequence 452, App	c 975	15	1.3	1491	4	US-09-285-479-171	Sequence 171, App
c 903	15	1.3	971	3	US-09-535-008-39	Sequence 39, Appl	976	15	1.3	1494	4	US-09-902-540-8469	Sequence 8469, Ap


```
; LENGTH: 2680
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4520

Query Match      1.6%; Score 19; DB 4; Length 2680;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 AGATGGAAAAAAGCCAAA 26
      |||||
Db      1963 AGATGGAAAAAAGCCAAA 1981

RESULT 5
US-09-620-312D-307
; Sequence 307, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: No. 6569662el Nucleic Acids and
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_Genes Version 1.0
; SEQ ID NO 307_
; LENGTH: 2982
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (198)..(2474)
US-09-620-312D-307

Query Match      1.6%; Score 19; DB 4; Length 2982;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 AGATGGAAAAAAGCCAAA 26
      |||||
Db      2248 AGATGGAAAAAAGCCAAA 2266

RESULT 6
US-09-949-016-15210/c
; Sequence 15210, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15210
; LENGTH: 5828
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15210

Query Match      1.6%; Score 19; DB 4; Length 5828;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 AGATGGAAAAAAGCCAAA 26
      |||||
Db      4249 AGATGGAAAAAAGCCAAA 4231

RESULT 7
US-09-949-016-16262
; Sequence 16262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16262
; LENGTH: 45571
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16262

Query Match      1.6%; Score 19; DB 4; Length 45571;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 AGATGGAAAAAAGCCAAA 26
      |||||
Db      41015 AGATGGAAAAAAGCCAAA 41033

RESULT 8
US-09-949-016-11942/c
; Sequence 11942, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
```

```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11942
; LENGTH: 66988
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(66988)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11942
```

```
Query Match 1.6%; Score 19; DB 4; Length 66988;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 109 ATTTAAACAATGTGATAAA 127
Db 44996 ATTTAAACAATGTGATAAA 44978
```

```
RESULT 9
US-09-949-016-16063/c
; Sequence 16063, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16063
; LENGTH: 66989
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(66989)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16063
```

```
Query Match 1.6%; Score 19; DB 4; Length 66989;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 109 ATTTAAACAATGTGATAAA 127
Db 44996 ATTTAAACAATGTGATAAA 44978
```

```
RESULT 10
US-09-949-016-16328
; Sequence 16328, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16328
; LENGTH: 68702
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16328
```

```
Query Match 1.6%; Score 19; DB 4; Length 68702;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 740 ATTGCTGCTGATGCAGCC 758
Db 66889 ATTGCTGCTGATGCAGCC 66907
```

```
RESULT 11
US-09-949-016-15617/c
; Sequence 15617, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15617
; LENGTH: 135171
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15617
```

```
Query Match 1.6%; Score 19; DB 4; Length 135171;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1160 AATTTACTTTTCTCTG 1178
Db 34518 AATTTACTTTTCTCTG 34500
```

```
RESULT 12
US-09-270-767-11620/c
; Sequence 11620, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11620
; LENGTH: 531
; TYPE: DNA
```

; ORGANISM: Drosophila melanogaster
US-09-270-767-11620

Query Match 1.5%; Score 18; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 634 GCCAGCGGAATGATTTC 651
|||||
Db 419 GCCAGCGGAATGATTTC 402

RESULT 13

US-09-949-016-55105
; Sequence 55105, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55105
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-55105

Query Match 1.5%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 893 CACAACAATGAATCACA 910
|||||
Db 353 CACAACAATGAATCACA 370

RESULT 14

US-09-949-016-55106
; Sequence 55106, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55106
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-55106

Query Match 1.5%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 52;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 893 CACAACAATGAATCACA 910
|||||
Db 452 CACAACAATGAATCACA 469

RESULT 15

US-09-949-016-76772
; Sequence 76772, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76772
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-76772

Query Match 1.5%; Score 18; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 888 ATAGGCACACAAATGAAA 905
|||||
Db 206 ATAGGCACACAAATGAAA 223

Search completed: June 4, 2005, 18:35:02
Job time : 298 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 15:33:57 ; Search time 802 Seconds
(without alignments)
9052.501 Million cells

Title: US-09-674-277-2

Perfect score: 1181

Sequence: 1 ctgcaggagatggaaaaaa.....ttttacttttttctctgcag 1181

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 5706582 seqs, 3073711274 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Minimum DB seq length: 9
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

```

Database :
Published Applications NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	21	1.8	341	9	US-09-983-965-1524	Sequence 1524, Ap	
C	2	21	1.8	578	18	US-10-425-115-160787	Sequence 160787,
	3	2	1.7	765	13	US-10-027-632-146278	Sequence 146278,
	4	2	1.7	765	13	US-10-027-632-146279	Sequence 146279,
	5	2	1.7	765	17	US-10-027-632-146278	Sequence 146278,
C	6	20	1.7	765	17	US-10-027-632-146279	Sequence 146279,
	7	20	1.7	1077	17	US-10-398-221-3166	Sequence 3166, Ap
	8	20	1.7	19016	19	US-10-469-442-1	Sequence 1, Appli
	9	20	1.7	319630	17	US-10-398-221-7	Sequence 7, Appli
C	10	20	1.7	3011208	17	US-10-398-221-2058	Sequence 2058, Ap
11	19	1.6	262	11	US-09-864-408A-3243	Sequence 3243, Ap	

c 85	18	1.5	1716	9	US-09-764-868-91	Sequence 91, Appl	c 158	17	1.4	509	9	US-09-989-279-149	Sequence 149, App
c 86	18	1.5	1773	15	US-10-081-872-155	Sequence 155, App	c 159	17	1.4	509	9	US-09-989-727-149	Sequence 149, App
c 87	18	1.5	1773	17	US-10-385-305-155	Sequence 155, App	c 160	17	1.4	509	9	US-09-989-731-149	Sequence 149, App
c 88	18	1.5	1856	18	US-10-425-115-96892	Sequence 96892, A	c 161	17	1.4	509	9	US-09-989-732-149	Sequence 149, App
c 89	18	1.5	2170	9	US-09-070-927A-543	Sequence 543, App	c 162	17	1.4	509	9	US-09-991-073-149	Sequence 149, App
c 90	18	1.5	2244	13	US-10-027-632-103667	Sequence 103667, A	c 163	17	1.4	509	9	US-09-990-442-149	Sequence 149, App
c 91	18	1.5	2244	17	US-10-027-632-103667	Sequence 103667, A	c 164	17	1.4	509	9	US-09-991-163-149	Sequence 149, App
c 92	18	1.5	2382	18	US-10-602-494-324	Sequence 324, App	c 165	17	1.4	509	9	US-09-993-604-149	Sequence 149, App
c 93	18	1.5	2427	17	US-10-369-493-24527	Sequence 24527, A	c 166	17	1.4	509	9	US-09-990-456-149	Sequence 149, App
c 94	18	1.5	2804	17	US-10-425-114-24869	Sequence 24869, A	c 167	17	1.4	509	9	US-09-989-721-149	Sequence 149, App
c 95	18	1.5	2894	15	US-10-037-270-95	Sequence 95, Appl	c 168	17	1.4	509	9	US-09-992-598-149	Sequence 149, App
c 96	18	1.5	2894	17	US-10-117-722-95	Sequence 95, Appl	c 169	17	1.4	509	9	US-09-989-293A-149	Sequence 149, App
c 97	18	1.5	2974	9	US-09-070-927A-422	Sequence 422, App	c 170	17	1.4	509	9	US-09-989-735-149	Sequence 149, App
c 98	18	1.5	3114	17	US-10-282-122A-38939	Sequence 38939, A	c 171	17	1.4	509	9	US-09-990-444-149	Sequence 149, App
c 99	18	1.5	3394	18	US-10-425-115-165568	Sequence 165568, A	c 172	17	1.4	509	9	US-09-991-181-149	Sequence 149, App
c 100	18	1.5	3606	19	US-10-717-665-37	Sequence 37, Appl	c 173	17	1.4	509	9	US-09-989-730-149	Sequence 149, App
c 101	18	1.5	3686	17	US-10-108-260A-1679	Sequence 1679, Ap	c 174	17	1.4	509	9	US-09-990-436-149	Sequence 149, App
c 102	18	1.5	4935	19	US-10-479-76A-43	Sequence 43, Appl	c 175	17	1.4	509	9	US-09-993-687-149	Sequence 149, App
c 103	18	1.5	26451	13	US-10-087-192-487	Sequence 487, App	c 176	17	1.4	509	10	US-09-989-734-149	Sequence 149, App
c 104	18	1.5	50000	14	US-10-152-722A-22	Sequence 22, Appl	c 177	17	1.4	509	10	US-09-997-653-149	Sequence 149, App
c 105	18	1.5	64183	18	US-10-684-422-201	Sequence 201, App	c 178	17	1.4	509	10	US-09-989-728-149	Sequence 149, App
c 106	18	1.5	107304	18	US-10-322-281-249	Sequence 249, App	c 179	17	1.4	509	10	US-09-989-728-149	Sequence 149, App
c 107	18	1.5	130877	18	US-10-322-281-54	Sequence 54, Appl	c 180	17	1.4	509	10	US-09-990-441-149	Sequence 149, App
c 108	18	1.5	165554	13	US-10-087-192-391	Sequence 391, App	c 181	17	1.4	509	10	US-09-993-667-149	Sequence 149, App
c 109	18	1.5	198073	13	US-10-087-192-682	Sequence 682, App	c 182	17	1.4	509	10	US-09-997-428-149	Sequence 149, App
c 110	18	1.5	260803	18	US-10-388-838-84	Sequence 84, Appl	c 183	17	1.4	509	10	US-09-997-666-149	Sequence 149, App
c 111	18	1.5	293625	18	US-10-719-993-6929	Sequence 6929, Ap	c 184	17	1.4	509	10	US-09-990-438-149	Sequence 149, App
c 112	18	1.5	379652	19	US-10-481-613-71	Sequence 71, Appl	c 185	17	1.4	509	10	US-09-990-562-149	Sequence 149, App
c 113	18	1.5	775062	18	US-10-719-993-6844	Sequence 6844, Ap	c 186	17	1.4	509	10	US-09-990-711-149	Sequence 149, App
c 114	18	1.5	783062	19	US-10-461-863-166	Sequence 166, App	c 187	17	1.4	509	10	US-09-989-726-149	Sequence 149, App
c 115	18	1.5	1830121	17	US-10-329-670-1	Sequence 1, Appl	c 188	17	1.4	509	10	US-09-998-156-149	Sequence 149, App
c 116	18	1.5	1830121	18	US-10-158-865-1	Sequence 1, Appl	c 189	17	1.4	509	10	US-09-990-437-149	Sequence 149, App
c 117	17	1.4	142	17	US-10-424-599-125819	Sequence 125819, A	c 190	17	1.4	509	10	US-09-991-157-149	Sequence 149, App
c 118	17	1.4	198	18	US-10-425-115-164684	Sequence 164684, A	c 191	17	1.4	509	10	US-09-997-514-149	Sequence 149, App
c 119	17	1.4	201	18	US-10-719-993-23896	Sequence 23896, A	c 192	17	1.4	509	10	US-09-997-573-149	Sequence 149, App
c 120	17	1.4	201	19	US-10-741-600-59377	Sequence 59377, A	c 193	17	1.4	509	10	US-09-991-172-149	Sequence 149, App
c 121	17	1.4	201	19	US-10-741-600-72145	Sequence 72145, A	c 194	17	1.4	509	10	US-09-990-726-149	Sequence 149, App
c 122	17	1.4	201	19	US-10-741-600-72148	Sequence 72148, A	c 195	17	1.4	509	10	US-09-997-559-149	Sequence 149, App
c 123	17	1.4	235	18	US-10-425-115-168456	Sequence 168456, A	c 196	17	1.4	509	10	US-09-997-601-149	Sequence 149, App
c 124	17	1.4	293	17	US-10-424-599-83508	Sequence 83508, A	c 197	17	1.4	509	10	US-09-990-443-149	Sequence 149, App
c 125	17	1.4	317	18	US-10-425-115-171294	Sequence 171294, A	c 198	17	1.4	509	10	US-09-991-854-149	Sequence 149, App
c 126	17	1.4	341	9	US-09-783-590-8145	Sequence 8145, Ap	c 199	17	1.4	509	10	US-09-997-628-149	Sequence 149, App
c 127	17	1.4	349	9	US-09-770-791-720	Sequence 720, App	c 200	17	1.4	509	10	US-09-997-683-149	Sequence 149, App
c 128	17	1.4	350	14	US-10-161-803-20	Sequence 20, Appl	c 201	17	1.4	509	10	US-09-989-729A-149	Sequence 149, App
c 129	17	1.4	353	9	US-09-764-847-1797	Sequence 1797, Ap	c 202	17	1.4	509	10	US-09-997-349-149	Sequence 149, App
c 130	17	1.4	353	9	US-09-764-847-1798	Sequence 1798, Ap	c 203	17	1.4	509	10	US-09-997-440-149	Sequence 149, App
c 131	17	1.4	353	14	US-10-092-154-1797	Sequence 1797, Ap	c 204	17	1.4	509	10	US-09-990-440-149	Sequence 149, App
c 132	17	1.4	368	17	US-10-092-154-1798	Sequence 1798, Ap	c 205	17	1.4	509	10	US-09-997-857-149	Sequence 149, App
c 133	17	1.4	368	17	US-10-424-599-94146	Sequence 94146, A	c 206	17	1.4	509	10	US-09-993-469-149	Sequence 149, App
c 134	17	1.4	369	17	US-10-424-599-34368	Sequence 34368, A	c 207	17	1.4	509	10	US-09-997-542-149	Sequence 149, App
c 135	17	1.4	380	9	US-09-933-797-364	Sequence 364, App	c 208	17	1.4	509	10	US-09-993-748-149	Sequence 149, App
c 136	17	1.4	381	18	US-10-425-115-85829	Sequence 85829, A	c 209	17	1.4	509	10	US-09-990-439-149	Sequence 149, App
c 137	17	1.4	391	18	US-10-357-930-7826	Sequence 7826, Ap	c 210	17	1.4	509	10	US-09-990-427-149	Sequence 149, App
c 138	17	1.4	414	18	US-10-674-124A-26338	Sequence 26338, A	c 211	17	1.4	509	10	US-09-989-328-149	Sequence 149, App
c 139	17	1.4	455	13	US-10-027-632-235818	Sequence 235818, A	c 212	17	1.4	509	10	US-09-993-583-149	Sequence 149, App
c 140	17	1.4	455	17	US-10-027-632-235818	Sequence 235818, A	c 213	17	1.4	509	10	US-09-991-992-149	Sequence 149, App
c 141	17	1.4	465	13	US-10-027-632-273502	Sequence 273502, A	c 214	17	1.4	509	10	US-09-992-521-149	Sequence 149, App
c 142	17	1.4	465	17	US-10-027-632-273502	Sequence 273502, A	c 215	17	1.4	509	10	US-09-997-333-149	Sequence 149, App
c 143	17	1.4	472	13	US-10-027-632-106720	Sequence 106720, A	c 216	17	1.4	509	10	US-09-997-384-149	Sequence 149, App
c 144	17	1.4	472	17	US-10-027-632-106720	Sequence 106720, A	c 217	17	1.4	509	10	US-09-998-041-149	Sequence 149, App
c 145	17	1.4	477	9	US-09-783-590-1232	Sequence 1232, Ap	c 218	17	1.4	509	10	US-09-997-585-149	Sequence 149, App
c 146	17	1.4	477	17	US-10-424-599-66801	Sequence 66801, A	c 219	17	1.4	509	10	US-09-997-614-149	Sequence 149, App
c 147	17	1.4	493	14	US-10-052-283-38	Sequence 38, Appl	c 220	17	1.4	509	10	US-09-989-862-149	Sequence 149, App
c 148	17	1.4	500	13	US-10-027-632-48525	Sequence 48525, A	c 221	17	1.4	509	10	US-09-997-529-149	Sequence 149, App
c 149	17	1.4	500	13	US-10-027-632-48526	Sequence 48526, A	c 222	17	1.4	509	10	US-09-989-725-149	Sequence 149, App
c 150	17	1.4	500	13	US-10-027-632-48527	Sequence 48527, A	c 223	17	1.4	509	10	US-09-991-150-149	Sequence 149, App
c 151	17	1.4	500	17	US-10-027-632-48525	Sequence 48525, A	c 224	17	1.4	509	10	US-09-997-641-149	Sequence 149, App
c 152	17	1.4	500	17	US-10-027-632-48526	Sequence 48526, A	c 225	17	1.4	509	10	US-09-989-733-149	Sequence 149, App
c 153	17	1.4	504	17	US-10-027-632-48527	Sequence 48527, A	c 226	17	1.4	509	10	US-09-992-643-149	Sequence 149, App
c 154	17	1.4	504	19	US-10-764-420-1095	Sequence 1095, Ap	c 227	17	1.4	509	17	US-10-219-538-149	Sequence 149, App
c 155	17	1.4	507	18	US-10-723-860-3854	Sequence 3854, Ap	c 228	17	1.4	520	9	US-09-927-850-8	Sequence 8, Appl
c 156	17	1.4	509	9	US-09-989-722-149	Sequence 149, App	c 229	17	1.4	520	18	US-10-021-323-1266	Sequence 13266, A
c 157	17	1.4	509	9	US-09-989-722-149	Sequence 149, App	c 230	17	1.4	534	13	US-10-021-323-1266	Sequence 95175, A

231	17	1.4	534	13	US-10-027-632-318627	Sequence 318627, A	304	17	1.4	913	17	US-10-027-632-163758	Sequence 163758, A
232	17	1.4	534	17	US-10-027-632-95175	Sequence 95175, A	305	17	1.4	925	18	US-10-437-963-48839	Sequence 48839, A
233	17	1.4	534	17	US-10-027-632-318627	Sequence 318627, A	306	17	1.4	981	9	US-09-822-849A-204	Sequence 204, App
234	17	1.4	534	18	US-10-425-115-114523	Sequence 114523, A	307	17	1.4	987	17	US-10-282-122A-11080	Sequence 11080, A
235	17	1.4	550	10	US-09-918-995-29534	Sequence 29534, A	308	17	1.4	995	17	US-10-425-114-8992	Sequence 8992, App
236	17	1.4	553	13	US-10-027-632-298721	Sequence 298721, A	309	17	1.4	1001	17	US-10-294-934-444	Sequence 444, App
237	17	1.4	553	17	US-10-027-632-298721	Sequence 298721, A	310	17	1.4	1005	17	US-10-425-114-18551	Sequence 18551, A
238	17	1.4	557	14	US-10-161-803-21	Sequence 21, Appl	311	17	1.4	1061	13	US-10-027-632-259258	Sequence 259258, A
239	17	1.4	567	18	US-10-437-963-19910	Sequence 19910, A	312	17	1.4	1061	17	US-10-027-632-259258	Sequence 259258, A
240	17	1.4	572	13	US-10-027-632-241342	Sequence 241342, A	313	17	1.4	1094	17	US-10-398-221-1866	Sequence 1866, App
241	17	1.4	572	17	US-10-027-632-241342	Sequence 241342, A	314	17	1.4	1111	18	US-10-767-701-10312	Sequence 10312, A
242	17	1.4	579	13	US-10-027-632-241452	Sequence 241452, A	315	17	1.4	1135	18	US-10-437-963-79922	Sequence 79922, A
243	17	1.4	579	13	US-10-027-632-224153	Sequence 224153, A	316	17	1.4	1141	18	US-10-437-963-79376	Sequence 79376, A
244	17	1.4	579	13	US-10-027-632-224154	Sequence 224154, A	317	17	1.4	1159	13	US-10-027-632-123325	Sequence 123325, A
245	17	1.4	579	13	US-10-027-632-224155	Sequence 224155, A	318	17	1.4	1159	17	US-10-027-632-123325	Sequence 123325, A
246	17	1.4	579	17	US-10-027-632-224152	Sequence 224152, A	319	17	1.4	1188	15	US-10-156-761-6944	Sequence 6944, App
247	17	1.4	579	17	US-10-027-632-224153	Sequence 224153, A	320	17	1.4	1193	9	US-09-919-344-17	Sequence 17, Appl
248	17	1.4	579	17	US-10-027-632-224154	Sequence 224154, A	321	17	1.4	1245	17	US-10-369-493-37240	Sequence 37240, A
249	17	1.4	579	17	US-10-027-632-224155	Sequence 224155, A	322	17	1.4	1245	17	US-10-027-632-123406	Sequence 123406, A
250	17	1.4	584	18	US-10-767-701-31041	Sequence 31041, A	323	17	1.4	1272	17	US-10-027-632-123406	Sequence 123406, A
251	17	1.4	592	18	US-10-357-930-53946	Sequence 53946, A	324	17	1.4	1325	17	US-10-282-122A-36453	Sequence 36453, A
252	17	1.4	595	18	US-10-425-115-161638	Sequence 161638, A	325	17	1.4	1335	16	US-10-380-913-1	Sequence 1, Appl
253	17	1.4	614	13	US-10-027-632-175082	Sequence 175082, A	326	17	1.4	1342	17	US-10-424-599-140285	Sequence 140285, A
254	17	1.4	614	17	US-10-027-632-175082	Sequence 175082, A	327	17	1.4	1376	13	US-10-027-632-251318	Sequence 251318, A
255	17	1.4	614	18	US-10-425-115-153338	Sequence 153338, A	328	17	1.4	1376	17	US-10-027-632-251318	Sequence 251318, A
256	17	1.4	626	18	US-10-767-701-25751	Sequence 25751, A	329	17	1.4	1376	17	US-10-424-599-137179	Sequence 137179, A
257	17	1.4	636	13	US-10-027-632-24225	Sequence 24225, A	330	17	1.4	1443	17	US-10-369-493-34301	Sequence 34301, A
258	17	1.4	636	17	US-10-027-632-24225	Sequence 24225, A	331	17	1.4	1446	15	US-10-220-382-23	Sequence 23, Appl
259	17	1.4	637	13	US-10-027-632-242595	Sequence 242595, A	332	17	1.4	1457	17	US-10-424-599-26058	Sequence 26058, A
260	17	1.4	637	17	US-10-027-632-242595	Sequence 242595, A	333	17	1.4	1524	17	US-10-424-599-50535	Sequence 50535, A
261	17	1.4	646	17	US-10-027-632-242595	Sequence 242595, A	334	17	1.4	1524	18	US-10-437-963-49816	Sequence 49816, A
262	17	1.4	652	11	US-09-969-034-4349	Sequence 4349, App	335	17	1.4	1536	18	US-10-437-963-5668	Sequence 5668, App
263	17	1.4	656	13	US-10-027-632-111116	Sequence 111116, A	336	17	1.4	1602	17	US-10-369-493-42846	Sequence 42846, A
264	17	1.4	656	17	US-10-027-632-111116	Sequence 111116, A	337	17	1.4	1708	15	US-10-037-270-275	Sequence 275, App
265	17	1.4	663	13	US-10-027-632-218070	Sequence 218070, A	338	17	1.4	1708	17	US-10-117-722-275	Sequence 275, App
266	17	1.4	663	17	US-10-027-632-218070	Sequence 218070, A	339	17	1.4	1715	17	US-10-276-774-806	Sequence 806, App
267	17	1.4	672	13	US-10-027-632-21878	Sequence 21878, A	340	17	1.4	1722	10	US-09-412-297-2	Sequence 2, Appl
268	17	1.4	672	13	US-10-027-632-21879	Sequence 21879, A	341	17	1.4	1815	14	US-10-107-643-9	Sequence 1, Appl
269	17	1.4	672	13	US-10-027-632-21880	Sequence 21880, A	342	17	1.4	1818	17	US-10-424-599-588	Sequence 588, App
270	17	1.4	672	17	US-10-027-632-21878	Sequence 21878, A	343	17	1.4	1902	17	US-10-398-221-1054	Sequence 1054, App
271	17	1.4	672	17	US-10-027-632-21879	Sequence 21879, A	344	17	1.4	1902	17	US-10-398-221-2831	Sequence 2831, App
272	17	1.4	672	17	US-10-027-632-21880	Sequence 21880, A	345	17	1.4	1905	18	US-10-437-963-31628	Sequence 31628, A
273	17	1.4	678	13	US-10-027-632-165929	Sequence 165929, A	346	17	1.4	1949	18	US-10-425-115-151706	Sequence 151706, A
274	17	1.4	678	17	US-10-027-632-165929	Sequence 165929, A	347	17	1.4	1950	17	US-10-152-319A-2100	Sequence 2100, App
275	17	1.4	680	13	US-10-027-632-111271	Sequence 111271, A	348	17	1.4	2068	17	US-10-108-260A-1459	Sequence 1459, App
276	17	1.4	680	13	US-10-027-632-111271	Sequence 111271, A	349	17	1.4	2181	17	US-10-282-122A-38893	Sequence 38893, A
277	17	1.4	680	17	US-10-027-632-111271	Sequence 111271, A	350	17	1.4	2181	17	US-10-282-122A-39763	Sequence 39763, A
278	17	1.4	680	17	US-10-027-632-111272	Sequence 111272, A	351	17	1.4	2274	19	US-10-741-849-6268	Sequence 6268, App
279	17	1.4	771	18	US-10-437-963-31140	Sequence 31140, A	352	17	1.4	2300	13	US-10-027-632-102778	Sequence 102778, A
280	17	1.4	775	13	US-10-027-632-168634	Sequence 168634, A	353	17	1.4	2300	17	US-10-027-632-102778	Sequence 102778, A
281	17	1.4	775	13	US-10-027-632-168635	Sequence 168635, A	354	17	1.4	2346	9	US-09-925-301-228	Sequence 228, App
282	17	1.4	775	17	US-10-027-632-168634	Sequence 168634, A	355	17	1.4	2362	17	US-10-258-662-91	Sequence 91, Appl
283	17	1.4	775	17	US-10-027-632-168635	Sequence 168635, A	356	17	1.4	2391	17	US-10-282-122A-11339	Sequence 11339, A
284	17	1.4	800	13	US-10-027-632-148591	Sequence 148591, A	357	17	1.4	2397	18	US-10-425-115-107544	Sequence 107544, A
285	17	1.4	800	17	US-10-027-632-148591	Sequence 148591, A	358	17	1.4	2403	17	US-10-369-493-42725	Sequence 42725, A
286	17	1.4	804	13	US-10-027-632-168952	Sequence 168952, A	359	17	1.4	2430	9	US-09-976-165-35	Sequence 35, Appl
287	17	1.4	804	13	US-10-027-632-168953	Sequence 168953, A	360	17	1.4	2430	13	US-10-027-632-110435	Sequence 110435, A
288	17	1.4	804	17	US-10-027-632-168952	Sequence 168952, A	361	17	1.4	2430	15	US-10-342-276-35	Sequence 35, Appl
289	17	1.4	804	17	US-10-027-632-168953	Sequence 168953, A	362	17	1.4	2646	16	US-10-027-632-110435	Sequence 110435, A
290	17	1.4	822	13	US-10-194-163-450	Sequence 450, App	363	17	1.4	2646	16	US-10-012-952A-96	Sequence 96, Appl
291	17	1.4	835	13	US-10-027-632-159647	Sequence 159647, A	364	17	1.4	2910	10	US-09-738-269-32	Sequence 32, Appl
292	17	1.4	835	13	US-10-027-632-159648	Sequence 159648, A	365	17	1.4	2910	13	US-10-023-437-32	Sequence 32, Appl
293	17	1.4	835	17	US-10-027-632-159647	Sequence 159647, A	366	17	1.4	2977	9	US-09-976-165-36	Sequence 36, Appl
294	17	1.4	835	17	US-10-027-632-159648	Sequence 159648, A	367	17	1.4	2977	9	US-09-954-456-735	Sequence 735, App
295	17	1.4	838	13	US-10-027-632-155981	Sequence 155981, A	368	17	1.4	2977	10	US-09-412-297-1	Sequence 1, Appl
296	17	1.4	838	17	US-10-027-632-155981	Sequence 155981, A	369	17	1.4	2977	15	US-10-342-276-36	Sequence 36, Appl
297	17	1.4	859	9	US-09-770-445-601	Sequence 601, App	370	17	1.4	2977	18	US-10-723-860-4019	Sequence 4019, App
298	17	1.4	913	9	US-09-927-850-1	Sequence 1, Appl	371	17	1.4	2977	19	US-10-843-641A-3762	Sequence 3762, App
299	17	1.4	913	13	US-10-027-632-163756	Sequence 163756, A	372	17	1.4	3096	17	US-10-607-631-3	Sequence 3, Appl
300	17	1.4	913	13	US-10-027-632-163757	Sequence 163757, A	373	17	1.4	3555	13	US-10-027-632-76197	Sequence 76197, A
301	17	1.4	913	17	US-10-027-632-163758	Sequence 163758, A	374	17	1.4	3555	17	US-10-027-632-76197	Sequence 76197, A
302	17	1.4	913	17	US-10-027-632-163756	Sequence 163756, A	375	17	1.4	3553	17	US-10-398-221-3748	Sequence 3748, App
303	17	1.4	913	17	US-10-027-632-163757	Sequence 163757, A	376	17	1.4	3816	13	US-10-027-632-76196	Sequence 76196, A

c 377	17	1.4	3816	17	US-10-027-632-76196	Sequence 76196, A	450	17	1.4	326014	9	US-09-731-231A-3	Sequence 3, Appli
378	17	1.4	3940	16	US-10-252-157-16	Sequence 16, Appl	451	17	1.4	326014	18	US-10-751-985-3	Sequence 3, Appli
379	17	1.4	6118	15	US-10-311-455-1004	Sequence 1004, Ap	452	17	1.4	385330	19	US-10-741-600-17796	Sequence 17796, A
c 380	17	1.4	6666	17	US-10-424-599-18523	Sequence 18523, A	c 453	17	1.4	385330	19	US-10-741-600-17796	Sequence 17796, A
381	17	1.4	7075	8	US-08-781-986A-263	Sequence 263, App	c 454	17	1.4	410846	19	US-10-481-613-1	Sequence 1, Appli
382	17	1.4	7075	17	US-10-329-624-263	Sequence 263, App	c 455	17	1.4	421609	18	US-10-367-094-122	Sequence 122, App
c 383	17	1.4	7315	17	US-10-424-599-18520	Sequence 18520, A	456	17	1.4	536165	10	US-09-939-964-1	Sequence 1, Appli
c 384	17	1.4	7737	9	US-09-764-887-453	Sequence 453, App	457	17	1.4	561515	18	US-10-741-601-5682	Sequence 5682, Ap
c 385	17	1.4	7737	9	US-09-764-847-2001	Sequence 2001, Ap	458	17	1.4	561515	19	US-10-741-600-17730	Sequence 17730, A
c 386	17	1.4	7737	14	US-10-092-154-2001	Sequence 2001, Ap	459	17	1.4	717651	18	US-10-719-993-6817	Sequence 6817, Ap
c 387	17	1.4	7737	14	US-10-073-961-453	Sequence 453, App	460	17	1.4	744802	17	US-10-292-798-1369	Sequence 1369, Ap
c 388	17	1.4	12258	17	US-10-307-817-137	Sequence 137, App	c 461	17	1.4	1830121	17	US-10-329-670-1	Sequence 1, Appli
389	17	1.4	14111	14	US-10-270-333-43	Sequence 43, Appl	c 462	17	1.4	1830121	18	US-10-158-865-1	Sequence 1, Appli
c 390	17	1.4	15479	17	US-10-257-166-46	Sequence 46, Appl	463	17	1.4	2731748	13	US-10-027-465A-1	Sequence 1, Appli
391	17	1.4	18705	11	US-09-984-429-463	Sequence 463, App	464	17	1.4	2940917	13	US-10-027-632-174763	Sequence 174763, A
392	17	1.4	18715	11	US-09-984-429-387	Sequence 387, App	465	17	1.4	2940917	17	US-10-027-632-174763	Sequence 174763, A
393	17	1.4	26656	16	US-10-229-834A-5	Sequence 5, Appli	c 466	17	1.4	3186778	13	US-10-027-632-174961	Sequence 174961, A
394	17	1.4	26750	13	US-10-087-192-1645	Sequence 1645, Ap	c 467	17	1.4	3186778	13	US-10-027-632-174961	Sequence 174961, A
c 395	17	1.4	28668	17	US-10-264-237-2863	Sequence 2863, Ap	468	17	1.4	3673778	16	US-10-312-841-2	Sequence 2, Appli
c 396	17	1.4	28668	17	US-10-264-237-2863	Sequence 2864, Ap	469	17	1.4	3673778	16	US-10-312-841-2	Sequence 2, Appli
397	17	1.4	32216	10	US-09-764-891-9613	Sequence 9613, Ap	470	17	1.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
398	17	1.4	32216	15	US-10-205-428-536	Sequence 536, App	471	16	1.4	18	16	US-10-133-779-64	Sequence 64, Appl
c 399	17	1.4	34063	14	US-10-114-170-96	Sequence 96, Appl	472	16	1.4	18	17	US-10-271-602B-104	Sequence 104, Appl
c 400	17	1.4	34185	15	US-10-199-520-3	Sequence 3, Appli	473	16	1.4	25	15	US-10-098-263B-71707	Sequence 71707, A
c 401	17	1.4	34185	15	US-10-874-827-3	Sequence 3, Appli	474	16	1.4	25	19	US-10-719-900-53221	Sequence 53221, A
402	17	1.4	39703	15	US-10-017-161-1041	Sequence 1041, Ap	475	16	1.4	25	19	US-10-719-900-315558	Sequence 315558, A
403	17	1.4	39703	15	US-10-292-798-883	Sequence 883, App	476	16	1.4	25	19	US-10-719-900-436614	Sequence 436614, A
c 404	17	1.4	45175	14	US-10-114-170-116	Sequence 116, App	c 476	16	1.4	25	19	US-10-719-900-505351	Sequence 505351, A
c 405	17	1.4	48509	18	US-10-322-281-510	Sequence 510, App	477	16	1.4	25	19	US-10-719-900-594172	Sequence 594172, A
c 406	17	1.4	49175	18	US-10-367-094-4	Sequence 4, Appl	c 478	16	1.4	25	19	US-10-719-900-809923	Sequence 809923, A
c 407	17	1.4	60940	17	US-10-052-482-68	Sequence 88, Appl	479	16	1.4	35	18	US-10-701-122-29	Sequence 29, Appl
c 408	17	1.4	70780	18	US-10-719-993-31819	Sequence 6819, Ap	480	16	1.4	65	10	US-09-908-975-28122	Sequence 28122, A
c 409	17	1.4	81001	9	US-09-751-877-1	Sequence 1, Appli	c 482	16	1.4	127	17	US-10-424-599-81692	Sequence 81692, A
c 410	17	1.4	81001	10	US-09-842-364-1	Sequence 1, Appli	c 483	16	1.4	166	13	US-10-040-739-917	Sequence 917, App
c 411	17	1.4	81001	11	US-09-751-877-1	Sequence 1, Appli	c 484	16	1.4	167	18	US-10-437-963-13826	Sequence 13826, A
c 412	17	1.4	81001	17	US-10-121-034-1	Sequence 1, Appli	c 485	16	1.4	186	17	US-10-242-535A-14914	Sequence 14914, A
c 413	17	1.4	81001	18	US-10-121-034-1	Sequence 1, Appli	c 486	16	1.4	186	17	US-10-085-783A-14914	Sequence 14914, A
c 414	17	1.4	83009	18	US-10-417-375-143	Sequence 143, App	c 487	16	1.4	186	18	US-10-425-115-158955	Sequence 158955, A
c 415	17	1.4	83405	19	US-10-723-518-1	Sequence 1, Appli	c 488	16	1.4	201	18	US-10-741-601-5942	Sequence 5942, Ap
c 416	17	1.4	94001	17	US-10-210-838-20	Sequence 20, Appl	c 489	16	1.4	201	18	US-10-741-601-21629	Sequence 21629, A
417	17	1.4	96960	19	US-10-484-577-662	Sequence 662, App	c 490	16	1.4	201	18	US-10-719-993-34877	Sequence 34877, A
c 418	17	1.4	102145	18	US-10-331-053-53	Sequence 53, Appl	c 491	16	1.4	201	18	US-10-719-993-34940	Sequence 34940, A
c 419	17	1.4	112132	9	US-09-741-150-3	Sequence 3, Appli	c 492	16	1.4	201	18	US-10-719-993-34987	Sequence 34987, A
c 420	17	1.4	112132	18	US-10-612-012-3	Sequence 3, Appli	c 493	16	1.4	201	18	US-10-719-993-41291	Sequence 41291, A
421	17	1.4	137197	10	US-09-754-853A-1	Sequence 1, Appli	c 494	16	1.4	201	18	US-10-719-993-42989	Sequence 42989, A
c 422	17	1.4	134292	17	US-10-240-425-1102	Sequence 1102, Ap	c 495	16	1.4	201	19	US-10-719-993-42989	Sequence 42989, A
c 423	17	1.4	134292	19	US-10-278-698-335	Sequence 335, App	c 496	16	1.4	201	19	US-10-741-600-18547	Sequence 18547, A
c 424	17	1.4	134292	19	US-10-278-698-849	Sequence 849, App	c 497	16	1.4	201	19	US-10-741-600-33559	Sequence 33559, A
c 425	17	1.4	135800	13	US-10-087-192-898	Sequence 898, App	c 498	16	1.4	201	19	US-10-741-600-33583	Sequence 33583, A
426	17	1.4	139389	17	US-10-236-031B-61	Sequence 61, Appl	499	16	1.4	210	17	US-10-242-535A-44821	Sequence 44821, A
c 427	17	1.4	154394	18	US-10-723-860-601	Sequence 601, App	500	16	1.4	210	17	US-10-085-783A-44821	Sequence 44821, A
c 428	17	1.4	162025	9	US-09-834-700-13	Sequence 13, Appl	c 501	16	1.4	217	18	US-10-674-124A-22163	Sequence 22163, A
c 429	17	1.4	162025	9	US-09-834-700-14	Sequence 14, Appl	c 502	16	1.4	231	18	US-10-674-124A-14820	Sequence 14820, A
c 430	17	1.4	162025	9	US-09-834-700-17	Sequence 17, Appl	c 503	16	1.4	233	16	US-10-029-386-23909	Sequence 23909, A
c 431	17	1.4	162025	9	US-09-834-700-17	Sequence 17, Appl	504	16	1.4	237	9	US-09-864-761-31042	Sequence 31042, A
c 432	17	1.4	162025	16	US-10-272-665-35	Sequence 35, Appl	c 505	16	1.4	247	9	US-09-563-817-864	Sequence 864, App
c 433	17	1.4	162025	16	US-10-272-665-36	Sequence 36, Appl	c 506	16	1.4	249	9	US-09-815-242-6765	Sequence 6765, Ap
c 434	17	1.4	162025	16	US-10-273-321-35	Sequence 35, Appl	c 507	16	1.4	270	9	US-09-815-242-3965	Sequence 3965, Ap
c 435	17	1.4	162025	16	US-10-273-321-36	Sequence 36, Appl	508	16	1.4	272	9	US-09-878-178-1960	Sequence 1960, Ap
c 436	17	1.4	162025	16	US-10-272-756-35	Sequence 35, Appl	509	16	1.4	272	13	US-10-046-935-1360	Sequence 1360, Ap
c 437	17	1.4	162025	16	US-10-272-756-36	Sequence 36, Appl	510	16	1.4	272	14	US-10-146-505-1360	Sequence 1360, Ap
c 438	17	1.4	162025	17	US-10-273-228-36	Sequence 35, Appl	c 511	16	1.4	278	15	US-10-102-524-403	Sequence 403, App
c 439	17	1.4	162025	17	US-10-273-228-36	Sequence 36, Appl	512	16	1.4	280	17	US-10-062-674-2151	Sequence 2151, Ap
440	17	1.4	168325	13	US-10-087-192-955	Sequence 955, App	513	16	1.4	282	18	US-10-425-115-101275	Sequence 101275, A
c 441	17	1.4	168749	17	US-10-085-117-250	Sequence 250, App	c 514	16	1.4	289	18	US-10-437-963-52735	Sequence 52735, A
c 442	17	1.4	175590	10	US-09-911-077A-13	Sequence 13, Appl	515	16	1.4	292	17	US-10-424-598-81651	Sequence 81651, A
c 443	17	1.4	175590	18	US-10-724-806-13	Sequence 13, Appl	c 516	16	1.4	295	17	US-10-424-598-112503	Sequence 112503, A
c 444	17	1.4	188971	17	US-10-235-192A-27	Sequence 27, Appl	c 517	16	1.4	295	18	US-10-425-115-106264	Sequence 106264, A
c 445	17	1.4	260027	13	US-10-087-192-298	Sequence 298, App	c 518	16	1.4	297	18	US-10-425-115-116779	Sequence 116779, A
446	17	1.4	261638	18	US-10-719-993-6882	Sequence 6882, Ap	c 519	16	1.4	298	17	US-10-125-968-793	Sequence 793, App
c 447	17	1.4	27616	18	US-10-367-094-83	Sequence 83, Appl	c 520	16	1.4	303	9	US-09-954-456-34	Sequence 34, Appl
c 448	17	1.4	289190	18	US-10-322-281-115	Sequence 115, App	c 521	16	1.4	303	9	US-09-954-456-990	Sequence 990, App
c 449	17	1.4	295096	13	US-10-087-192-331	Sequence 331, App	c 522	16	1.4	303	19	US-10-843-641A-3061	Sequence 3061, Ap

c 523	16	1.4	303	19	US-10-843-641A-4017	Sequence 4017, Ap	596	16	1.4	421	17	US-10-027-632-313027	Sequence 313027,
c 524	16	1.4	305	18	US-10-425-115-138814	Sequence 138814, Ap	c 597	16	1.4	421	17	US-10-424-599-127981	Sequence 127981,
c 525	16	1.4	306	10	US-09-764-891-6924	Sequence 6924, Ap	c 598	16	1.4	422	9	US-09-969-708-4373	Sequence 437, App
c 526	16	1.4	307	16	US-10-253-990-1	Sequence 1, Appl	599	16	1.4	422	9	US-09-783-590-6384	Sequence 6384, Ap
c 527	16	1.4	308	10	US-09-991-936-309	Sequence 309, App	600	16	1.4	422	10	US-09-764-891-1960	Sequence 1960, Ap
c 528	16	1.4	309	10	US-09-764-891-6923	Sequence 6923, Ap	c 601	16	1.4	422	19	US-10-843-641A-7904	Sequence 7904, Ap
c 529	16	1.4	310	17	US-10-621-901-369	Sequence 369, App	602	16	1.4	426	9	US-09-983-965-429	Sequence 429, App
c 530	16	1.4	313	10	US-09-764-891-800	Sequence 800, App	603	16	1.4	426	18	US-10-021-323-4556	Sequence 4556, Ap
c 531	16	1.4	315	18	US-10-437-963-57473	Sequence 57473, A	c 604	16	1.4	429	10	US-09-918-995-4908	Sequence 4908, Ap
c 532	16	1.4	316	17	US-10-125-968-790	Sequence 790, App	c 605	16	1.4	440	17	US-10-424-599-24451	Sequence 24451, A
c 533	16	1.4	317	15	US-10-254-676-41	Sequence 41, Appl	c 606	16	1.4	444	17	US-10-242-535A-36748	Sequence 36748, A
c 534	16	1.4	317	17	US-10-125-968-784	Sequence 784, App	c 607	16	1.4	444	17	US-10-085-783A-36748	Sequence 36748, A
c 535	16	1.4	317	17	US-10-125-968-825	Sequence 825, App	c 608	16	1.4	448	9	US-09-864-761-4838	Sequence 4838, Ap
c 536	16	1.4	320	17	US-10-125-968-824	Sequence 824, App	c 609	16	1.4	448	18	US-10-674-124A-11384	Sequence 11384, A
c 537	16	1.4	324	9	US-09-974-300-7250	Sequence 7250, App	c 610	16	1.4	450	18	US-10-324-967-50	Sequence 50, Appl
c 538	16	1.4	325	13	US-10-027-632-71163	Sequence 71163, A	611	16	1.4	450	10	US-09-918-995-15892	Sequence 15892, A
c 539	16	1.4	325	17	US-10-027-632-71163	Sequence 71163, A	612	16	1.4	457	18	US-10-437-963-90943	Sequence 90943, A
c 540	16	1.4	325	18	US-10-027-632-71163	Sequence 71163, A	c 613	16	1.4	461	9	US-09-864-761-15166	Sequence 15166, A
c 541	16	1.4	330	13	US-10-079-623-56	Sequence 56, Appl	614	16	1.4	463	10	US-09-814-353-16952	Sequence 16952, A
c 542	16	1.4	335	9	US-09-867-701-1986	Sequence 1986, App	615	16	1.4	465	17	US-10-242-535A-51964	Sequence 51964, A
c 543	16	1.4	336	9	US-09-864-761-23955	Sequence 23955, A	616	16	1.4	465	17	US-10-085-783A-51964	Sequence 51964, A
c 544	16	1.4	339	10	US-09-940-727B-99	Sequence 99, Appl	c 617	16	1.4	468	10	US-09-918-995-32648	Sequence 32648, A
c 545	16	1.4	339	17	US-10-125-968-822	Sequence 822, App	618	16	1.4	468	13	US-10-027-632-191621	Sequence 191621, A
c 546	16	1.4	342	17	US-10-424-599-141651	Sequence 141651, A	619	16	1.4	468	17	US-10-027-632-191621	Sequence 191621, A
c 547	16	1.4	346	14	US-10-198-846-12240	Sequence 12240, A	620	16	1.4	469	17	US-10-242-535A-42852	Sequence 42852, A
c 548	16	1.4	349	17	US-10-125-968-794	Sequence 794, App	621	16	1.4	469	17	US-10-085-783A-42852	Sequence 42852, A
c 549	16	1.4	350	9	US-09-920-300A-989	Sequence 989, App	c 622	16	1.4	469	18	US-10-425-115-94502	Sequence 94502, A
c 550	16	1.4	350	13	US-10-033-528-989	Sequence 989, App	c 623	16	1.4	470	17	US-10-125-968-884	Sequence 884, App
c 551	16	1.4	350	16	US-10-099-926-989	Sequence 989, App	624	16	1.4	471	17	US-10-424-599-129784	Sequence 129784, A
c 552	16	1.4	351	17	US-10-125-968-823	Sequence 823, App	c 625	16	1.4	475	11	US-09-732-627A-4491	Sequence 4491, Ap
c 553	16	1.4	353	17	US-10-424-599-65910	Sequence 65910, A	c 626	16	1.4	478	14	US-10-198-846-11860	Sequence 11860, A
c 554	16	1.4	353	18	US-10-430-201-3976	Sequence 3976, App	c 627	16	1.4	480	17	US-10-424-599-104803	Sequence 104803, A
c 555	16	1.4	353	18	US-10-430-201-3977	Sequence 3977, App	628	16	1.4	480	10	US-09-770-961-954	Sequence 954, App
c 556	16	1.4	357	17	US-10-242-535A-53806	Sequence 53806, A	629	16	1.4	482	17	US-10-424-599-102669	Sequence 102669, A
c 557	16	1.4	357	18	US-10-085-783A-53806	Sequence 53806, A	c 630	16	1.4	485	9	US-09-867-701-29511	Sequence 29511, Ap
c 558	16	1.4	357	18	US-10-425-115-115364	Sequence 115364, A	631	16	1.4	485	10	US-09-918-995-28641	Sequence 28641, A
c 559	16	1.4	359	10	US-09-803-719-2012	Sequence 2012, Ap	c 632	16	1.4	487	13	US-10-027-632-195164	Sequence 195164, A
c 560	16	1.4	364	18	US-10-674-124A-12251	Sequence 12251, A	c 633	16	1.4	487	17	US-10-027-632-195164	Sequence 195164, A
c 561	16	1.4	364	18	US-10-425-115-145349	Sequence 145349, A	c 634	16	1.4	489	9	US-09-864-761-2796	Sequence 2796, Ap
c 562	16	1.4	367	17	US-10-242-535A-23211	Sequence 23211, A	c 635	16	1.4	493	9	US-09-783-590-4234	Sequence 4234, Ap
c 563	16	1.4	367	17	US-10-085-783A-23211	Sequence 23211, A	c 636	16	1.4	494	10	US-09-918-995-22409	Sequence 22409, A
c 564	16	1.4	368	10	US-09-940-727B-103	Sequence 103, App	c 637	16	1.4	494	11	US-09-969-034-3864	Sequence 3864, Ap
c 565	16	1.4	368	10	US-09-940-727B-107	Sequence 107, App	c 638	16	1.4	494	13	US-10-027-632-306233	Sequence 306233, A
c 566	16	1.4	374	17	US-10-424-599-124300	Sequence 124300, A	c 639	16	1.4	494	17	US-10-027-632-306233	Sequence 306233, A
c 567	16	1.4	375	17	US-10-424-599-97078	Sequence 97078, A	c 640	16	1.4	496	17	US-10-424-599-25138	Sequence 25138, A
c 568	16	1.4	377	18	US-10-437-963-46675	Sequence 46675, A	c 641	16	1.4	498	10	US-09-918-995-14626	Sequence 14626, A
c 569	16	1.4	379	9	US-09-796-692-3163	Sequence 3163, Ap	642	16	1.4	499	17	US-10-424-599-20082	Sequence 20082, A
c 570	16	1.4	379	14	US-10-040-862-3163	Sequence 3163, Ap	643	16	1.4	501	16	US-10-029-386-9808	Sequence 9808, Ap
c 571	16	1.4	379	17	US-10-057-475B-3163	Sequence 3163, Ap	644	16	1.4	501	16	US-10-029-386-10197	Sequence 10197, A
c 572	16	1.4	379	17	US-10-154-884B-3163	Sequence 3163, Ap	c 645	16	1.4	501	17	US-10-424-599-73422	Sequence 73422, A
c 573	16	1.4	379	18	US-10-764-324-3163	Sequence 3163, Ap	c 646	16	1.4	502	13	US-10-027-632-96506	Sequence 96506, A
c 574	16	1.4	384	18	US-10-425-115-67509	Sequence 67509, A	c 647	16	1.4	502	17	US-10-027-632-96506	Sequence 96506, A
c 575	16	1.4	387	17	US-10-424-599-50899	Sequence 50899, A	648	16	1.4	503	10	US-09-814-353-21139	Sequence 21139, A
c 576	16	1.4	396	18	US-10-425-115-156323	Sequence 156323, A	649	16	1.4	503	14	US-10-198-846-12858	Sequence 12858, A
c 577	16	1.4	399	9	US-09-954-531-839	Sequence 531, App	c 650	16	1.4	507	10	US-09-918-995-15209	Sequence 15209, A
c 578	16	1.4	399	9	US-09-954-531-839	Sequence 531, App	c 651	16	1.4	508	18	US-10-425-115-94569	Sequence 94569, A
c 579	16	1.4	399	17	US-10-242-535A-10735	Sequence 10735, A	c 652	16	1.4	510	9	US-09-864-761-8474	Sequence 8474, Ap
c 580	16	1.4	399	17	US-10-085-783A-10735	Sequence 10735, A	653	16	1.4	510	16	US-10-029-386-4981	Sequence 4981, Ap
c 581	16	1.4	399	19	US-10-843-641A-1906	Sequence 1906, Ap	654	16	1.4	510	13	US-10-437-963-34210	Sequence 34210, A
c 582	16	1.4	399	19	US-10-843-641A-2325	Sequence 2325, Ap	655	16	1.4	513	18	US-10-027-632-70354	Sequence 70354, A
c 583	16	1.4	400	9	US-09-867-701-5344	Sequence 5344, Ap	656	16	1.4	513	17	US-10-027-632-70354	Sequence 70354, A
c 584	16	1.4	400	14	US-10-198-846-9569	Sequence 9569, Ap	c 657	16	1.4	513	17	US-10-403-571-71	Sequence 71, Appl
c 585	16	1.4	405	9	US-09-983-965-4387	Sequence 4387, Ap	658	16	1.4	514	17	US-10-424-599-86645	Sequence 86645, A
c 586	16	1.4	405	10	US-09-918-995-17836	Sequence 17836, A	659	16	1.4	514	13	US-10-027-632-24213	Sequence 24213, A
c 587	16	1.4	409	10	US-09-918-995-5898	Sequence 5898, Ap	660	16	1.4	516	13	US-10-027-632-24214	Sequence 24214, A
c 588	16	1.4	414	17	US-10-125-968-881	Sequence 881, App	661	16	1.4	516	17	US-10-027-632-24213	Sequence 24213, A
c 589	16	1.4	415	9	US-09-864-761-5727	Sequence 5727, Ap	662	16	1.4	516	17	US-10-027-632-24214	Sequence 24214, A
c 590	16	1.4	417	17	US-10-424-599-106459	Sequence 106459, A	c 663	16	1.4	518	13	US-10-027-632-191617	Sequence 191617, A
c 591	16	1.4	419	9	US-09-867-701-817	Sequence 817, App	c 664	16	1.4	518	13	US-10-027-632-191618	Sequence 191618, A
c 592	16	1.4	420	10	US-09-940-727B-111	Sequence 111, App	c 665	16	1.4	518	17	US-10-027-632-191617	Sequence 191617, A
c 593	16	1.4	421	13	US-10-027-632-73343	Sequence 73343, A	c 666	16	1.4	520	17	US-10-027-632-191618	Sequence 191618, A
c 594	16	1.4	421	13	US-10-027-632-733027	Sequence 733027, A	c 667	16	1.4	520	17	US-10-424-599-67949	Sequence 67949, A
c 595	16	1.4	421	17	US-10-027-632-73343	Sequence 73343, A	668	16	1.4	525	17	US-10-424-599-119691	Sequence 119691, A

c 669	16	1.4	526	18	US-10-425-115-152445	Sequence 152445,	742	16	1.4	586	13	US-10-027-632-92985	Sequence 92985, A
c 670	16	1.4	526	18	US-10-357-930-58440	Sequence 58440, A	743	16	1.4	586	13	US-10-027-632-275822	Sequence 275822,
c 671	16	1.4	528	13	US-10-027-632-195770	Sequence 195770,	744	16	1.4	586	17	US-10-027-632-92985	Sequence 92985, A
c 672	16	1.4	528	17	US-10-027-632-195770	Sequence 195770,	745	16	1.4	586	17	US-10-027-632-275822	Sequence 275822,
c 673	16	1.4	529	13	US-10-027-632-275157	Sequence 275157,	746	16	1.4	588	16	US-10-029-388-11910	Sequence 11910, A
c 674	16	1.4	529	17	US-10-027-632-275157	Sequence 275157,	c 747	16	1.4	588	18	US-10-437-963-81443	Sequence 81443, A
c 675	16	1.4	530	14	US-10-198-846-9044	Sequence 9044, Ap	748	16	1.4	589	13	US-10-027-632-86169	Sequence 86169, A
c 676	16	1.4	531	9	US-09-884-441-96	Sequence 96, Appl	749	16	1.4	589	17	US-10-027-632-86169	Sequence 86169, A
c 677	16	1.4	531	10	US-09-907-969-96	Sequence 96, Appl	750	16	1.4	590	9	US-09-920-300A-245	Sequence 245, App
c 678	16	1.4	531	10	US-09-827-271-96	Sequence 96, Appl	751	16	1.4	590	13	US-10-033-528-245	Sequence 245, App
c 679	16	1.4	531	15	US-10-198-053-96	Sequence 96, Appl	752	16	1.4	590	16	US-10-099-928-245	Sequence 245, App
c 680	16	1.4	531	19	US-10-860-790-96	Sequence 96, Appl	753	16	1.4	590	18	US-10-425-115-73681	Sequence 73681, A
c 681	16	1.4	533	19	US-10-425-115-171747	Sequence 171747,	c 754	16	1.4	592	10	US-09-764-891-7365	Sequence 7365, Ap
c 682	16	1.4	534	9	US-09-796-692-4781	Sequence 4781, Ap	755	16	1.4	592	17	US-10-259-194A-514	Sequence 514, App
c 683	16	1.4	534	14	US-10-040-862-4781	Sequence 4781, Ap	c 756	16	1.4	598	13	US-10-027-632-246529	Sequence 246529,
c 684	16	1.4	534	17	US-10-057-475B-4781	Sequence 4781, Ap	c 757	16	1.4	598	13	US-10-027-632-246530	Sequence 246530,
c 685	16	1.4	534	17	US-10-154-884B-4781	Sequence 4781, Ap	c 758	16	1.4	598	13	US-10-027-632-246531	Sequence 246531,
c 686	16	1.4	534	18	US-10-764-324-4781	Sequence 4781, Ap	c 759	16	1.4	598	17	US-10-027-632-246529	Sequence 246529,
c 687	16	1.4	535	13	US-10-027-632-275823	Sequence 275823,	c 760	16	1.4	598	17	US-10-027-632-246530	Sequence 246530,
c 688	16	1.4	535	17	US-10-027-632-275823	Sequence 275823,	c 761	16	1.4	598	17	US-10-027-632-246531	Sequence 246531,
c 689	16	1.4	536	13	US-10-027-632-231448	Sequence 231448,	c 762	16	1.4	599	18	US-10-425-115-165977	Sequence 165977,
c 690	16	1.4	536	13	US-10-027-632-231449	Sequence 231449,	c 763	16	1.4	600	18	US-10-357-930-54090	Sequence 54090, A
c 691	16	1.4	536	17	US-10-027-632-231448	Sequence 231448,	c 764	16	1.4	601	18	US-10-834-850-15	Sequence 15, Appl
c 692	16	1.4	537	16	US-10-027-632-231449	Sequence 231449,	c 765	16	1.4	601	18	US-10-255-536-98	Sequence 98, Appl
c 693	16	1.4	537	16	US-10-029-386-12253	Sequence 12253, A	c 766	16	1.4	603	14	US-10-437-963-5498	Sequence 5498, A
c 694	16	1.4	538	13	US-10-027-632-94884	Sequence 94884, A	c 767	16	1.4	608	13	US-10-027-632-65102	Sequence 65102, A
c 695	16	1.4	538	13	US-10-027-632-318557	Sequence 318557,	c 768	16	1.4	608	13	US-10-027-632-65102	Sequence 65102, A
c 696	16	1.4	538	17	US-10-027-632-94884	Sequence 94884, A	c 769	16	1.4	608	17	US-10-027-632-65102	Sequence 65102, A
c 697	16	1.4	538	17	US-10-027-632-318557	Sequence 318557,	c 770	16	1.4	608	17	US-10-027-632-65102	Sequence 65102, A
c 698	16	1.4	539	9	US-09-796-692-4730	Sequence 4730, Ap	c 771	16	1.4	608	18	US-10-021-323-5336	Sequence 5336, Ap
c 699	16	1.4	539	14	US-10-040-862-4730	Sequence 4730, Ap	c 772	16	1.4	608	18	US-10-425-115-122665	Sequence 122665,
c 700	16	1.4	539	17	US-10-057-475B-4730	Sequence 4730, Ap	c 773	16	1.4	609	13	US-10-027-632-108006	Sequence 108006,
c 701	16	1.4	539	17	US-10-154-884B-4730	Sequence 4730, Ap	c 774	16	1.4	609	17	US-10-027-632-108006	Sequence 108006,
c 702	16	1.4	539	18	US-10-764-324-4730	Sequence 4730, Ap	c 775	16	1.4	610	18	US-10-767-701-29790	Sequence 29790, A
c 703	16	1.4	539	18	US-10-425-115-137818	Sequence 137818,	c 776	16	1.4	612	13	US-10-027-632-206208	Sequence 206208,
c 704	16	1.4	539	19	US-10-487-901-5992	Sequence 5992, Ap	c 777	16	1.4	612	13	US-10-027-632-206209	Sequence 206209,
c 705	16	1.4	540	17	US-10-424-598-126274	Sequence 126274,	c 778	16	1.4	612	17	US-10-027-632-206208	Sequence 206208,
c 706	16	1.4	542	18	US-10-437-963-99248	Sequence 99248, A	c 779	16	1.4	612	17	US-10-027-632-206209	Sequence 206209,
c 707	16	1.4	546	13	US-10-027-632-199787	Sequence 199787,	c 780	16	1.4	613	13	US-10-027-632-106868	Sequence 106868,
c 708	16	1.4	546	13	US-10-027-632-199788	Sequence 199788,	c 781	16	1.4	613	13	US-10-027-632-106869	Sequence 106869,
c 709	16	1.4	546	13	US-10-027-632-199789	Sequence 199789,	c 782	16	1.4	613	17	US-10-027-632-106868	Sequence 106868,
c 710	16	1.4	546	17	US-10-027-632-199787	Sequence 199787,	c 783	16	1.4	613	17	US-10-027-632-106869	Sequence 106869,
c 711	16	1.4	546	17	US-10-027-632-199788	Sequence 199788,	c 784	16	1.4	618	13	US-10-027-632-208480	Sequence 208480,
c 712	16	1.4	546	17	US-10-027-632-199789	Sequence 199789,	c 785	16	1.4	618	17	US-10-027-632-208480	Sequence 208480,
c 713	16	1.4	548	9	US-09-864-761-6609	Sequence 6609, Ap	c 786	16	1.4	621	18	US-10-468-488-448	Sequence 448, App
c 714	16	1.4	552	9	US-09-864-761-7224	Sequence 7224, Ap	c 787	16	1.4	621	18	US-10-437-963-30539	Sequence 30539, A
c 715	16	1.4	554	18	US-10-425-115-67011	Sequence 67011, A	c 788	16	1.4	623	17	US-10-424-593-61199	Sequence 61199, A
c 716	16	1.4	553	13	US-10-027-632-129151	Sequence 129151,	c 789	16	1.4	624	13	US-10-027-632-241230	Sequence 241230,
c 717	16	1.4	563	13	US-10-027-632-129152	Sequence 129152,	c 790	16	1.4	624	17	US-10-027-632-241230	Sequence 241230,
c 718	16	1.4	563	17	US-10-027-632-129151	Sequence 129151,	c 791	16	1.4	624	17	US-10-424-599-61084	Sequence 61084, A
c 719	16	1.4	563	17	US-10-027-632-129152	Sequence 129152,	c 792	16	1.4	625	18	US-10-425-115-142312	Sequence 142312,
c 720	16	1.4	565	18	US-10-767-701-26163	Sequence 26163, A	c 793	16	1.4	627	18	US-10-437-963-58428	Sequence 58428, A
c 721	16	1.4	568	18	US-10-021-323-5399	Sequence 5399, Ap	c 794	16	1.4	628	13	US-10-027-632-47528	Sequence 47528, A
c 722	16	1.4	574	13	US-10-027-632-320714	Sequence 320714,	c 795	16	1.4	630	17	US-10-027-632-47528	Sequence 47528, A
c 723	16	1.4	574	13	US-10-027-632-320715	Sequence 320715,	c 796	16	1.4	630	17	US-10-424-599-5876	Sequence 5876, Ap
c 724	16	1.4	574	17	US-10-027-632-320715	Sequence 320715,	c 797	16	1.4	631	18	US-10-767-701-6617	Sequence 6617, Ap
c 725	16	1.4	574	17	US-10-027-632-320715	Sequence 320715,	c 798	16	1.4	635	13	US-10-027-632-202095	Sequence 202095,
c 726	16	1.4	574	18	US-10-425-115-134124	Sequence 134124,	c 799	16	1.4	635	13	US-10-027-632-202096	Sequence 202096,
c 727	16	1.4	575	13	US-10-027-632-257544	Sequence 257544,	c 800	16	1.4	635	17	US-10-027-632-202095	Sequence 202095,
c 728	16	1.4	575	13	US-10-027-632-257545	Sequence 257545,	c 801	16	1.4	635	17	US-10-027-632-202096	Sequence 202096,
c 729	16	1.4	575	17	US-10-027-632-257544	Sequence 257544,	c 802	16	1.4	638	17	US-10-425-114-15312	Sequence 15312, A
c 730	16	1.4	575	17	US-10-027-632-257545	Sequence 257545,	c 803	16	1.4	641	18	US-10-425-115-132977	Sequence 132977,
c 731	16	1.4	578	16	US-10-027-632-10003	Sequence 10003, A	c 804	16	1.4	644	13	US-10-027-632-214832	Sequence 214832,
c 732	16	1.4	580	13	US-10-027-632-172467	Sequence 172467,	c 805	16	1.4	644	17	US-10-027-632-214832	Sequence 214832,
c 733	16	1.4	580	13	US-10-027-632-222849	Sequence 222849,	c 806	16	1.4	645	13	US-10-027-632-86252	Sequence 86252, A
c 734	16	1.4	580	17	US-10-027-632-172467	Sequence 172467,	c 807	16	1.4	645	13	US-10-027-632-86253	Sequence 86253, A
c 735	16	1.4	580	17	US-10-027-632-222849	Sequence 222849,	c 808	16	1.4	645	17	US-10-027-632-86252	Sequence 86252, A
c 736	16	1.4	585	13	US-10-027-632-319580	Sequence 319580,	c 809	16	1.4	645	17	US-10-027-632-86253	Sequence 86253, A
c 737	16	1.4	585	13	US-10-027-632-319581	Sequence 319581,	c 810	16	1.4	646	13	US-10-027-632-41494	Sequence 41494, A
c 738	16	1.4	585	13	US-10-027-632-319582	Sequence 319582,	c 811	16	1.4	646	13	US-10-027-632-41495	Sequence 41495, A
c 739	16	1.4	585	17	US-10-027-632-319580	Sequence 319580,	c 812	16	1.4	646	13	US-10-027-632-69770	Sequence 69770, A
c 740	16	1.4	585	17	US-10-027-632-319581	Sequence 319581,	c 813	16	1.4	646	13	US-10-027-632-69771	Sequence 69771, A
c 741	16	1.4	585	17	US-10-027-632-319582	Sequence 319582,	c 814	16	1.4	646	13	US-10-027-632-70365	Sequence 70365, A

815	16	1.4	646	17	US-10-027-632-41494	Sequence 41494, A	888	16	1.4	747	17	US-10-027-632-173039	Sequence 173039,
816	16	1.4	646	17	US-10-027-632-41495	Sequence 41495, A	889	16	1.4	747	17	US-10-027-632-173040	Sequence 173040,
817	16	1.4	646	17	US-10-027-632-69770	Sequence 69770, A	890	16	1.4	756	17	US-10-282-122A-25922	Sequence 25922, A
818	16	1.4	646	17	US-10-027-632-69771	Sequence 69771, A	891	16	1.4	759	10	US-09-814-353-4376	Sequence 4276, A
819	16	1.4	646	17	US-10-027-632-70365	Sequence 70365, A	892	16	1.4	759	10	US-09-814-353-10580	Sequence 10580, A
820	16	1.4	647	13	US-10-027-632-70593	Sequence 70593, A	893	16	1.4	759	17	US-10-369-493-46608	Sequence 46608, A
821	16	1.4	647	13	US-10-027-632-312206	Sequence 312206, A	894	16	1.4	759	18	US-10-37-963-19380	Sequence 19380, A
822	16	1.4	647	17	US-10-027-632-70593	Sequence 70593, A	895	16	1.4	762	16	US-10-407-138-1	Sequence 1, Appl1
823	16	1.4	647	17	US-10-027-632-312206	Sequence 312206, A	896	16	1.4	765	17	US-10-424-599-109937	Sequence 109937, A
824	16	1.4	649	14	US-10-198-846-5701	Sequence 5701, Ap	897	16	1.4	770	18	US-10-425-115-49297	Sequence 49297, A
825	16	1.4	649	18	US-10-425-115-149643	Sequence 149643, A	898	16	1.4	771	14	US-10-198-846-4364	Sequence 4364, Ap
826	16	1.4	650	13	US-10-027-632-202291	Sequence 202291, A	899	16	1.4	771	17	US-10-424-599-50255	Sequence 50255, A
827	16	1.4	650	13	US-10-027-632-202292	Sequence 202292, A	900	16	1.4	786	14	US-10-198-846-4405	Sequence 4405, Ap
828	16	1.4	650	13	US-10-027-632-202293	Sequence 202293, A	901	16	1.4	789	13	US-10-027-632-139598	Sequence 139598, A
829	16	1.4	650	13	US-10-027-632-202294	Sequence 202294, A	902	16	1.4	789	17	US-10-027-632-139598	Sequence 139598, A
830	16	1.4	650	13	US-10-027-632-204217	Sequence 204217, A	903	16	1.4	789	18	US-10-437-963-29743	Sequence 29743, A
831	16	1.4	650	17	US-10-027-632-202291	Sequence 202291, A	904	16	1.4	792	13	US-10-027-632-7679	Sequence 7679, Ap
832	16	1.4	650	17	US-10-027-632-202292	Sequence 202292, A	905	16	1.4	792	14	US-10-198-846-4417	Sequence 4417, Ap
833	16	1.4	650	17	US-10-027-632-202293	Sequence 202293, A	906	16	1.4	792	17	US-10-027-632-7679	Sequence 7679, Ap
834	16	1.4	650	17	US-10-027-632-202294	Sequence 202294, A	907	16	1.4	794	17	US-10-369-493-38588	Sequence 38588, A
835	16	1.4	650	17	US-10-027-632-204217	Sequence 204217, A	908	16	1.4	801	9	US-09-967-772-7	Sequence 7, Appl1
836	16	1.4	654	13	US-10-027-632-46099	Sequence 46099, A	909	16	1.4	801	17	US-10-335-517-7	Sequence 7, Appl1
837	16	1.4	654	17	US-10-027-632-46099	Sequence 46099, A	910	16	1.4	801	17	US-10-334-288-7	Sequence 7, Appl1
838	16	1.4	654	17	US-10-027-632-25018	Sequence 25018, A	911	16	1.4	801	17	US-10-311-030-6	Sequence 6, Appl1
839	16	1.4	656	13	US-10-027-632-25019	Sequence 25019, A	912	16	1.4	801	18	US-10-425-115-114557	Sequence 114557, A
840	16	1.4	656	17	US-10-027-632-25018	Sequence 25018, A	913	16	1.4	801	18	US-10-656-023-23	Sequence 23, Appl1
841	16	1.4	656	17	US-10-027-632-25019	Sequence 25019, A	914	16	1.4	804	14	US-10-198-846-4433	Sequence 4433, Ap
842	16	1.4	663	13	US-10-027-632-213201	Sequence 213201, A	915	16	1.4	804	14	US-10-198-846-5651	Sequence 5651, Ap
843	16	1.4	663	17	US-10-027-632-213201	Sequence 213201, A	916	16	1.4	814	14	US-10-198-846-5651	Sequence 5651, Ap
844	16	1.4	665	14	US-10-198-846-4439	Sequence 4439, Ap	917	16	1.4	818	13	US-10-027-632-171918	Sequence 171918, A
845	16	1.4	665	18	US-10-425-115-59095	Sequence 59095, A	918	16	1.4	818	17	US-10-027-632-171918	Sequence 171918, A
846	16	1.4	669	14	US-10-198-846-4475	Sequence 4475, Ap	919	16	1.4	819	14	US-10-198-846-4434	Sequence 4434, Ap
847	16	1.4	672	9	US-09-822-849A-590	Sequence 590, App	920	16	1.4	819	14	US-10-198-846-4434	Sequence 4434, Ap
848	16	1.4	677	18	US-10-653-047-4656	Sequence 4656, Ap	921	16	1.4	823	14	US-10-198-846-5606	Sequence 5606, Ap
849	16	1.4	679	13	US-10-027-632-141999	Sequence 141999, A	922	16	1.4	826	14	US-10-198-846-5638	Sequence 5638, Ap
850	16	1.4	679	17	US-10-027-632-141999	Sequence 141999, A	923	16	1.4	830	14	US-10-198-846-5639	Sequence 5639, Ap
851	16	1.4	684	18	US-10-437-963-19381	Sequence 19381, A	924	16	1.4	833	14	US-10-198-846-7174	Sequence 7174, Ap
852	16	1.4	687	13	US-10-027-632-260998	Sequence 260998, A	925	16	1.4	842	17	US-10-424-599-106721	Sequence 106721, A
853	16	1.4	687	13	US-10-027-632-260998	Sequence 260998, A	926	16	1.4	848	14	US-10-198-846-5674	Sequence 5674, Ap
854	16	1.4	687	17	US-10-027-632-260998	Sequence 260998, A	927	16	1.4	848	14	US-10-198-846-7159	Sequence 7159, Ap
855	16	1.4	687	17	US-10-027-632-260999	Sequence 260999, A	928	16	1.4	850	14	US-10-198-846-7159	Sequence 7159, Ap
856	16	1.4	688	18	US-10-653-047-5332	Sequence 5332, Ap	929	16	1.4	851	13	US-10-027-632-136819	Sequence 126819, A
857	16	1.4	690	13	US-10-027-632-222562	Sequence 222562, A	930	16	1.4	851	13	US-10-027-632-136819	Sequence 126820, A
858	16	1.4	690	17	US-10-027-632-222562	Sequence 222562, A	931	16	1.4	851	13	US-10-027-632-136819	Sequence 126820, A
859	16	1.4	695	17	US-10-341-961A-23	Sequence 23, Appl1	932	16	1.4	851	17	US-10-027-632-136819	Sequence 126819, A
860	16	1.4	703	17	US-10-424-599-92113	Sequence 92113, A	933	16	1.4	851	17	US-10-027-632-136820	Sequence 126820, A
861	16	1.4	706	18	US-10-437-963-101091	Sequence 101091, A	934	16	1.4	851	17	US-10-027-632-153718	Sequence 153718, A
862	16	1.4	713	18	US-10-767-701-7247	Sequence 7247, Ap	935	16	1.4	854	13	US-10-027-632-2761	Sequence 2761, Ap
863	16	1.4	714	19	US-10-487-901-7059	Sequence 7059, Ap	936	16	1.4	854	13	US-10-027-632-2762	Sequence 2762, Ap
864	16	1.4	718	13	US-10-027-632-166763	Sequence 166763, A	937	16	1.4	854	13	US-10-027-632-2763	Sequence 2763, Ap
865	16	1.4	718	13	US-10-027-632-166764	Sequence 166764, A	938	16	1.4	854	17	US-10-027-632-2763	Sequence 2763, Ap
866	16	1.4	718	13	US-10-027-632-166765	Sequence 166765, A	939	16	1.4	854	17	US-10-027-632-2763	Sequence 2763, Ap
867	16	1.4	718	13	US-10-027-632-166766	Sequence 166766, A	940	16	1.4	854	17	US-10-027-632-2763	Sequence 2763, Ap
868	16	1.4	718	17	US-10-027-632-166763	Sequence 166763, A	941	16	1.4	860	17	US-10-296-115-342	Sequence 342, App
869	16	1.4	718	17	US-10-027-632-166764	Sequence 166764, A	942	16	1.4	863	17	US-10-424-599-41519	Sequence 41519, A
870	16	1.4	718	17	US-10-027-632-166765	Sequence 166765, A	943	16	1.4	870	14	US-10-198-846-5543	Sequence 5543, Ap
871	16	1.4	725	17	US-10-027-632-166766	Sequence 166766, A	944	16	1.4	872	13	US-10-027-632-166412	Sequence 166412, A
872	16	1.4	725	17	US-10-424-599-54898	Sequence 54898, A	945	16	1.4	872	13	US-10-027-632-166413	Sequence 166413, A
873	16	1.4	726	18	US-10-767-701-8709	Sequence 8709, Ap	946	16	1.4	872	17	US-10-027-632-166414	Sequence 166414, A
874	16	1.4	728	13	US-10-027-632-24215	Sequence 24215, A	947	16	1.4	872	17	US-10-027-632-166412	Sequence 166412, A
875	16	1.4	728	17	US-10-027-632-24215	Sequence 24215, A	948	16	1.4	872	17	US-10-027-632-166413	Sequence 166413, A
876	16	1.4	732	17	US-10-424-599-522	Sequence 522, App	949	16	1.4	872	17	US-10-027-632-166414	Sequence 166414, A
877	16	1.4	734	9	US-09-939-980-51	Sequence 51, Appl1	950	16	1.4	876	14	US-10-198-846-3540	Sequence 3540, Ap
878	16	1.4	740	13	US-10-027-632-18033	Sequence 18033, A	951	16	1.4	880	14	US-10-198-846-7140	Sequence 7140, Ap
879	16	1.4	740	13	US-10-027-632-18034	Sequence 18034, A	952	16	1.4	884	13	US-10-027-632-255514	Sequence 255514, A
880	16	1.4	740	17	US-10-027-632-18033	Sequence 18033, A	953	16	1.4	884	14	US-10-198-846-7040	Sequence 7040, Ap
881	16	1.4	740	17	US-10-027-632-18034	Sequence 18034, A	954	16	1.4	884	17	US-10-027-632-255514	Sequence 255514, A
882	16	1.4	743	13	US-10-016-634A-9	Sequence 9, Appl1	955	16	1.4	885	17	US-10-424-599-105813	Sequence 105813, A
883	16	1.4	743	18	US-10-425-115-109971	Sequence 109971, A	956	16	1.4	886	14	US-10-198-846-5717	Sequence 5717, Ap
884	16	1.4	747	13	US-10-027-632-173038	Sequence 173038, A	957	16	1.4	888	14	US-10-198-846-4467	Sequence 4467, Ap
885	16	1.4	747	13	US-10-027-632-173039	Sequence 173039, A	958	16	1.4	888	14	US-10-198-846-7131	Sequence 7131, Ap
886	16	1.4	747	13	US-10-027-632-173040	Sequence 173040, A	959	16	1.4	897	14	US-10-198-846-4326	Sequence 4326, Ap
887	16	1.4	747	17	US-10-027-632-173038	Sequence 173038, A	960	16	1.4	897	14	US-10-198-846-4454	Sequence 4454, Ap


```
961 1.4 898 18 US-10-437-963-45609 Sequence 45609, A
962 1.4 903 14 US-10-198-846-7039 Sequence 7039, Ap
963 1.4 919 14 US-10-198-846-4324 Sequence 4324, Ap
964 1.4 921 9 US-09-898-570-25 Sequence 25, Appl
965 1.4 921 10 US-09-839-446-25 Sequence 25, Appl
966 1.4 922 13 US-10-199-330-2 Sequence 2, Appl
967 1.4 922 14 US-10-199-334-2 Sequence 2, Appl
968 1.4 922 16 US-10-199-329-2 Sequence 2, Appl
969 1.4 925 13 US-10-027-632-28583 Sequence 28583, A
970 1.4 925 17 US-10-027-632-28583 Sequence 28583, A
971 1.4 926 17 US-10-369-493-37946 Sequence 37946, A
972 1.4 926 17 US-10-369-493-38262 Sequence 38262, A
973 1.4 929 14 US-10-198-846-7150 Sequence 7150, Ap
974 1.4 942 13 US-10-199-330-1 Sequence 1, Appl
975 1.4 942 14 US-10-199-334-1 Sequence 1, Appl
976 1.4 942 16 US-10-199-329-1 Sequence 1, Appl
977 1.4 954 18 US-10-425-115-96743 Sequence 96743, A
978 1.4 972 17 US-10-282-122A-37032 Sequence 37032, A
979 1.4 975 17 US-10-166-653-5 Sequence 5, Appl
980 1.4 975 17 US-10-152-319A-2118 Sequence 2118, Ap
981 1.4 978 17 US-10-282-122A-39078 Sequence 39078, A
982 1.4 978 17 US-10-282-122A-39863 Sequence 39863, A
983 1.4 991 18 US-10-767-701-9158 Sequence 9158, Ap
984 1.4 1004 13 US-10-199-330-3 Sequence 3, Appl
985 1.4 1004 14 US-10-199-334-3 Sequence 3, Appl
986 1.4 1004 16 US-10-199-329-3 Sequence 3, Appl
987 1.4 1009 14 US-10-198-846-5718 Sequence 5718, Ap
988 1.4 1044 18 US-10-437-963-18942 Sequence 18942, A
989 1.4 1045 19 US-10-487-078-130 Sequence 130, App
990 1.4 1046 13 US-10-027-632-118680 Sequence 118680,
991 1.4 1046 13 US-10-027-632-118681 Sequence 118681,
992 1.4 1046 17 US-10-027-632-118680 Sequence 118680,
993 1.4 1046 17 US-10-027-632-118681 Sequence 118681,
994 1.4 1050 17 US-10-282-122A-40502 Sequence 40502, A
995 1.4 1088 17 US-10-425-114-10080 Sequence 10080, A
996 1.4 1098 14 US-10-074-475-49 Sequence 49, Appl
997 1.4 1098 17 US-10-671-403-143 Sequence 143, App
998 1.4 1098 17 US-10-671-419-143 Sequence 143, App
999 1.4 1098 17 US-10-670-844-143 Sequence 143, App
c1000 1.4 1098 17 US-10-671-134-143 Sequence 143, App
```

ALIGNMENTS

```
RESULT 1
US-09-983-965-1524
; Sequence 1524, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathalagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 1524
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (297),(300)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 37-LIB2809-032-Q1-E1-B2
```

```
US-09-983-965-1524
Query Match 1.8%; Score 21; DB 9; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1077 AGTCTGAGGATGGAAGGAGG 1097
|||||
Db 245 AGTCTGAGGATGGAAGGAGG 265

RESULT 2
US-10-425-115-160787/c
; Sequence 160787, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 160787
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(578)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: MRT4577_78218C.1
US-10-425-115-160787

Query Match 1.8%; Score 21; DB 18; Length 578;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AAAAAAGCCAAATAAAAAA 34
|||||
Db 574 AAAAAAGCCAAATAAAAAA 554

RESULT 3
US-10-027-632-146278
; Sequence 146278, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
```



```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 146278
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-146278
```

```
Query Match 1.7%; Score 20; DB 13; Length 765;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1118 AATAAAGTGACATCATGCC 1137
|||||
Db 712 AATAAAGTGACATCATGCC 731
```

RESULT 4

```
US-10-027-632-146279
; Sequence 146279, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 146279
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-146279
```

```
Query Match 1.7%; Score 20; DB 13; Length 765;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1118 AATAAAGTGACATCATGCC 1137
|||||
Db 712 AATAAAGTGACATCATGCC 731
```

RESULT 5

```
US-10-027-632-146278
; Sequence 146278, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
```

```
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 146278
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-146278
```

```
Query Match 1.7%; Score 20; DB 17; Length 765;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1118 AATAAAGTGACATCATGCC 1137
|||||
Db 712 AATAAAGTGACATCATGCC 731
```

RESULT 6

```
US-10-027-632-146279
; Sequence 146279, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 146279
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-146279
```

```
Query Match 1.7%; Score 20; DB 17; Length 765;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1118 AATAAAGTGACATCATGCC 1137
|||||
Db 712 AATAAAGTGACATCATGCC 731
```

RESULT 7

```
US-10-398-221-3166/c
; Sequence 3166, Application US/10398221
; Publication No. US20040018514A1
```

; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3166
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3166

Query Match 1.7%; Score 20; DB 17; Length 1077;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 683 CATAAATTCATGTCCTTTT 702
|||||
Db 773 CATAAATTCATGTCCTTTT 754

RESULT 8
US-10-469-442-1
; Sequence 1, Application US/10469442
; Publication No. US20050089954A1
; GENERAL INFORMATION:
; APPLICANT: PALMU, Kaisa
; APPLICANT: KUNNARI, Teto
; TITLE OF INVENTION: GENE CLUSTER FOR RABELOMYCIN BIOSYNTHESIS, AND ITS USE TO GENERATE COMPOUNDS FOR DRUG SCREENING
; FILE REFERENCE: 0933-0208P
; CURRENT APPLICATION NUMBER: US/10/469,442
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: PCT/FI02/00214
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 22
; SEQ ID NO 1
; LENGTH: 19016
; TYPE: DNA
; ORGANISM: Streptomyces sp. H021
US-10-469-442-1

Query Match 1.7%; Score 20; DB 19; Length 19016;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 GCCCGGCGCCACACGAGC 635
|||||
Db 10504 GCCCGGCGCCACACGAGC 10523

RESULT 9
US-10-398-221-7
; Sequence 7, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27

; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 319630
; TYPE: DNA
; ORGANISM: Listeria innocua
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
US-10-398-221-7

Query Match 1.7%; Score 20; DB 17; Length 319630;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 683 CATAAATTCATGTCCTTTT 702
|||||
Db 289402 CATAAATTCATGTCCTTTT 289421

RESULT 10
US-10-398-221-2058/c
; Sequence 2058, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua
US-10-398-221-2058

Query Match 1.7%; Score 20; DB 17; Length 3011208;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 683 CATAAATTCATGTCCTTTT 702
|||||
Db 1828021 CATAAATTCATGTCCTTTT 1828002

RESULT 11
US-09-864-408A-3243
; Sequence 3243, Application US/09864408A
; Publication No. US20040009474A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Enc
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3243

; LENGTH: 262
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-408A-3243

Query Match 1.6%; Score 19; DB 11; Length 262;
Best Local Similarity 100.0%; Pred. No. 32; Mismatches 0; Indels 0; Gaps 0;

Qy 1076 AAGTCTCAGGATGGAAGGA 1094
|||||
Db 237 AAGTCTCAGGATGGAAGGA 255

RESULT 12

US-09-796-692-8712
; Sequence 8712, Application US/09796692
; Publication No. US20020198362A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.

; APPLICANT: Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY

; FILE REFERENCE: 2077.001200

; CURRENT APPLICATION NUMBER: US/09/796,692

; PRIOR FILING DATE: 2001-03-01

; PRIOR APPLICATION NUMBER: 60/186,126

; PRIOR FILING DATE: 2000-03-01

; PRIOR APPLICATION NUMBER: 60/190,479

; PRIOR FILING DATE: 2000-03-17

; PRIOR APPLICATION NUMBER: 60/200,545

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 60/200,303

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,779

; PRIOR FILING DATE: 2000-04-28

; PRIOR APPLICATION NUMBER: 60/200,999

; PRIOR FILING DATE: 2000-05-01

; PRIOR APPLICATION NUMBER: 60/202,084

; PRIOR FILING DATE: 2000-05-04

; PRIOR APPLICATION NUMBER: 60/206,201

; PRIOR FILING DATE: 2000-05-22

; PRIOR APPLICATION NUMBER: 60/218,950

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/222,903

; PRIOR FILING DATE: 2000-08-03

; PRIOR APPLICATION NUMBER: 60/223,416

; PRIOR FILING DATE: 2000-08-04

; PRIOR APPLICATION NUMBER: 60/223,378

; PRIOR FILING DATE: 2000-08-07

; SOFTWARE: FastSeq for Windows Version 3.0

; NUMBER OF SEQ ID NOS: 9597

; SEQ ID NO 8712

; LENGTH: 477

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-796-692-8712

Query Match 1.6%; Score 19; DB 9; Length 477;
Best Local Similarity 100.0%; Pred. No. 33; Mismatches 0; Indels 0; Gaps 0;

Qy 740 ATTGCTGCTGATGCAGCC 758
|||||
Db 231 ATTGCTGCTGATGCAGCC 249

RESULT 13

US-10-040-862-8712

; Sequence 8712, Application US/10040862

; Publication No. US20030078396A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8712
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-8712

Query Match 1.6%; Score 19; DB 14; Length 477;
Best Local Similarity 100.0%; Pred. No. 33; Mismatches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 740 ATTGCTGCTGATGCAGCC 758
|||||
Db 231 ATTGCTGCTGATGCAGCC 249

RESULT 14

US-10-057-475B-8712
; Sequence 8712, Application US/10057475B
; Publication No. US20040002068A1

; GENERAL INFORMATION:

; APPLICANT: Gaiger, Alexander

; APPLICANT: Algate, Paul A.

; APPLICANT: Mannion, Jane

; APPLICANT: Clapper, Jonathan David

; APPLICANT: Wang, Aijun

; APPLICANT: Ordonez, Nadia

; APPLICANT: Carter, Lauren

; APPLICANT: McNeill, Patricia Dianne

; APPLICANT: Corixa Corporation

; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therap

; FILE REFERENCE: 014058-014402US

; CURRENT APPLICATION NUMBER: US/10/057,475B

; CURRENT FILING DATE: 2002-01-22

; PRIOR APPLICATION NUMBER: US 60/186,126

```
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8712
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-057-475B-8712
```

```
Query Match 1.6%; Score 19; DB 17; Length 477;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 740 ATTGTCGTGCTGATGCAGCC 758
    |||||
Db 231 ATTGTCGTGCTGATGCAGCC 249
```

RESULT 15

```
US-10-154-884B-8712
; Sequence 8712, Application US/10154884B
; Publication No. US20040005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8712
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-8712

Query Match 1.6%; Score 19; DB 17; Length 477;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 740 ATTGTCGTGCTGATGCAGCC 758
    |||||
Db 231 ATTGTCGTGCTGATGCAGCC 249

Search completed: June 4, 2005, 18:48:15
Job time : 847 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 15:16:17 ; Search time 4425 Seconds
(without alignments)
10159.074 Million cell updates/sec

Title: US-09-674-277-2
Perfect score: 1181
Sequence: 1 ctccagagatgtaaaaaa.....ttttactttttctctgcag 1181

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : EST.*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gest1.*
- 9: gb_gest2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	22	1.9	635	2 BB657961	BB657961 BB657961
C 2	22	1.9	638	9 AG044971	AG044971 Pan trogl
C 3	22	1.9	650	2 BB469169	BB469169 BB469169
C 4	22	1.9	787	4 B1225178	B1225178 602949857
C 5	22	1.9	888	3 AK084319	AK084319 Mus muscu
C 6	21	1.8	174	2 BF612800	BF612800 dd76906.x
C 7	21	1.8	329	9 CL385624	CL385624 RPI44_33
C 8	21	1.8	390	2 BE478771	BE478771 163162 BA
C 9	21	1.8	416	1 AV668123	AV668123 AV668123
C 10	21	1.8	430	1 AV596235	AV596235 AV596235
C 11	21	1.8	441	8 AQ594469	AQ594469 HS_2105_B
C 12	21	1.8	529	8 AZ475198	AZ475198 1M0293J19
C 13	21	1.8	672	2 BB659888	BB659888 BB659888
C 14	21	1.8	717	1 AJ452665	AJ452665 AJ452665
C 15	21	1.8	727	6 BY761301	BY761301 BY761301
C 16	21	1.8	750	1 AJ449744	AJ449744 AJ449744
C 17	21	1.8	763	9 CG809836	CG809836 FSAAC71TF
C 18	21	1.8	778	1 AJ456064	AJ456064 AJ456064
C 19	21	1.8	837	7 CK747810	CK747810 963604 MA
C 20	21	1.8	940	5 BU128223	BU128223 603113931
C 21	20	1.7	232	2 BB214339	BB214339 BB214339
C 22	20	1.7	277	1 AV097971	AV097971 AV097971
C 23	20	1.7	279	2 BB581122	BB581122 BB581122
C 24	20	1.7	375	5 BP636732	BP636732 BP636732

CF287163	AGENCOURT	385	7	CF287163	20	1.7
AQ752982	HS_5226_B	410	8	AQ752982	20	1.7
EX036346	Single re	27		CNS09071	26	
BE678140	df62h04.x	491	3	BE678140	20	1.7
CD893441	G118.123K	541	6	CD893441	20	1.7
BI103911	602890168	554	6	BI103911	20	1.7
CD893440	G118.123K	631	6	CD893440	20	1.7
BD432815	BONIHA8TF	690	6	BD432815	20	1.7
BX208511	CH230-437	747	8	BX208511	20	1.7
BX854388	BX854388	795	5	BX854388	20	1.7
CG951752	MBRJC69TF	854	5	CG951752	20	1.7
CL132899	ISB1-102L	895	9	CL132899	20	1.7
CH098898	ke14d06.y	989	9	CH098898	20	1.7
CB122664	K-EST0170	157	6	CB122664	19	1.6
AJ589279	Arabidops	161	6	AJ589279	19	1.6
BB199280	BB199280	171	9	BB199280	19	1.6
AV358502	AV358502	232	1	AV358502	19	1.6
CR744154	CR744154	237	1	CR744154	19	1.6
AV143506	AV143506	249	7	AV143506	19	1.6
AV226877	AV226877	259	1	AV226877	19	1.6
BB054209	BB054209	289	1	BB054209	19	1.6
AZ698251	RPCR-23-2	278	2	AZ698251	19	1.6
AA249101	sh1117.be	305	8	AA249101	19	1.6
CL254850	ZMWB060	306	1	CL254850	19	1.6
BG727702	fo91h07.x	331	9	BG727702	19	1.6
CB507202	sealtc018	337	4	CB507202	19	1.6
AV687426	AV687426	342	6	AV687426	19	1.6
AW547187	LO018E07-	349	1	AW547187	19	1.6
AV654008	AV654008	355	2	AV654008	19	1.6
AQ044211	CIT-HSP-2	368	1	AQ044211	19	1.6
AQ193727	CIT-HSP-2	369	8	AQ193727	19	1.6
CD340145	StrPu336.	370	8	CD340145	19	1.6
AI428773	mp28b05.x	376	6	AI428773	19	1.6
CR526012	CR526012	378	1	CR526012	19	1.6
AG261403	Lotus cor	381	7	AG261403	19	1.6
CB099029	ko05c12.y	386	9	CB099029	19	1.6
BE179387	IL3-HT061	399	6	BE179387	19	1.6
CK381463	leg38g03.	405	2	CK381463	19	1.6
AI840382	UI-M-AH0-	418	7	AI840382	19	1.6
BE116779	UI-R-B81-	421	1	BE116779	19	1.6
CR131269	Reverse s	428	5	CR131269	19	1.6
CE486797	tigr-g88-	432	9	CE486797	19	1.6
CD740102	4028850.1	445	9	CD740102	19	1.6
EX482292	DKFZp866P	453	6	EX482292	19	1.6
CB512637	sealrqb54	456	5	CB512637	19	1.6
CE154362	tigr-g88-	462	6	CE154362	19	1.6
AG265330	Lotus cor	472	9	AG265330	19	1.6
AI283347	qui14e08.x	473	9	AI283347	19	1.6
CK746274	wm101.5mb	484	1	CK746274	19	1.6
BI044927	MR4-OT010	487	7	BI044927	19	1.6
H662299	yr72f04.r1	497	4	H662299	19	1.6
N92222	yz90605.r1	498	7	N92222	19	1.6
BM739599	K-EST0009	499	7	BM739599	19	1.6
AG253793	Lotus cor	515	4	AG253793	19	1.6
H60782	yr45all.r1	516	4	H60782	19	1.6
CL802211	OR_CBA001	517	7	CL802211	19	1.6
BM646032	170006873	518	9	BM646032	19	1.6
AL491057	T. brucei	522	4	AL491057	19	1.6
BF334191	RC1-CT024	524	9	BF334191	19	1.6
CK900493	SGP156420	528	2	CK900493	19	1.6
BH858538	BS_x037a	532	7	BH858538	19	1.6
AZ341294	EST006069	533	8	AZ341294	19	1.6
BQ496840	EST006069	534	8	BQ496840	19	1.6
AQ023556	CGR0472A	535	5	AQ023556	19	1.6
CV475898	24531.1 D	539	8	CV475898	19	1.6
AL464773	T. brucei	543	7	AL464773	19	1.6
BF110512	7n38d04.x	560	9	BF110512	19	1.6
CO739692	170006875	562	2	CO739692	19	1.6
BM636003	SL0006a24	565	7	BM636003	19	1.6
BP338402	BP338402	570	4	BP338402	19	1.6
BP576466	602133884	571	5	BP576466	19	1.6
BP368892	BP368892	578	5	BP368892	19	1.6

C 98	19	1.6	580	2	BE568467	BB568467	601342553	171	19	1.6	718	7	CK464559	CK464559	935573	MA
C 99	19	1.6	585	2	BF129824	BF129824	601817603	c 172	19	1.6	718	7	CK466649	CK466649	937877	MA
C 100	19	1.6	586	9	CL803623	CL803623	OR_CBA001	c 173	19	1.6	723	7	CK351267	CK351267	hgghfa35F	
C 101	19	1.6	587	4	BG616229	BG616229	602644734	c 174	19	1.6	724	4	BG532107	BG532107	602561157	
C 102	19	1.6	588	2	BF212447	BF212447	601813215	c 175	19	1.6	726	9	AG440186	AG440186	Mus muscu	
C 103	19	1.6	590	2	BE619529	BE619529	601473085	c 176	19	1.6	735	9	CL747759	CL747759	OR_BBA011	
C 104	19	1.6	590	4	BG502492	BG502492	602549104	c 177	19	1.6	736	9	CNS062BB	AL422217	T7 end of	
C 105	19	1.6	590	9	CG952747	CG952747	MBBHJ83TR	c 178	19	1.6	741	9	AG584325	AG584325	Mus muscu	
C 106	19	1.6	591	4	BG474184	BG474184	602516782	c 179	19	1.6	741	9	CE272879	CE272879	tigr-g88-	
C 107	19	1.6	594	8	BZ190393	BZ190393	CH230-279	c 180	19	1.6	744	6	CB999243	CB999243	AGENCOURT	
C 108	19	1.6	598	2	BF892794	BF892794	QV1-MT013	c 181	19	1.6	753	7	CO397061	CO397061	AGENCOURT	
C 109	19	1.6	598	2	BF129006	BF129006	601811327	c 182	19	1.6	754	7	CR370140	CR370140	CR370140	
C 110	19	1.6	600	2	BF028796	BF028796	601764755	c 183	19	1.6	757	1	AJ794612	AJ794612	AJ794612	
C 111	19	1.6	600	4	BG503204	BG503204	602551303	c 184	19	1.6	762	6	CA369499	CA369499	645961	NC
C 112	19	1.6	601	2	BF791185	BF791185	602251285	c 185	19	1.6	763	2	BE368455	BE368455	601220539	
C 113	19	1.6	601	2	BF132553	BF132553	601645411	c 186	19	1.6	772	2	BE541536	BE541536	601067807	
C 114	19	1.6	602	2	BE897464	BE897464	601439385	c 187	19	1.6	778	6	CA414814	CA414814	UI-H-E20-	
C 115	19	1.6	604	4	BG505558	BG505558	602549271	c 188	19	1.6	783	5	BU621484	BU621484	UI-H-E20-	
C 116	19	1.6	605	5	BX254985	BX254985	60254985	c 189	19	1.6	783	6	CA420232	CA420232	UI-H-E20-	
C 117	19	1.6	608	4	BG494916	BG494916	602540934	c 190	19	1.6	787	5	BU629467	BU629467	UI-H-FLO-	
C 118	19	1.6	609	4	BG117610	BG117610	6023347955	c 191	19	1.6	789	4	BG503915	BG503915	602552811	
C 119	19	1.6	610	9	CE000463	CE000463	tigr-g88-	c 192	19	1.6	790	8	CC137316	CC137316	NDL-91M21	
C 120	19	1.6	612	4	B1089399	B1089399	602854627	c 193	19	1.6	792	8	BZ449607	BZ449607	BONHB49TR	
C 121	19	1.6	616	4	B1868892	B1868892	603394968	c 194	19	1.6	795	4	BG494592	BG494592	602540646	
C 122	19	1.6	617	4	BM740426	BM740426	K-EST0011	c 195	19	1.6	796	2	BF128835	BF128835	601811020	
C 123	19	1.6	618	7	CN166707	CN166707	998387	c 196	19	1.6	799	6	CA368128	CA368128	644200	NC
C 124	19	1.6	619	4	BG613526	BG613526	602641459	c 197	19	1.6	804	5	BU961128	BU961128	AGENCOURT	
C 125	19	1.6	620	4	BG778657	BG778657	602667930	c 198	19	1.6	806	9	CR272393	CR272393	Reverse s	
C 126	19	1.6	624	9	CK341792	CK341792	Medicargo	c 199	19	1.6	807	5	BX867217	BX867217	601066378	
C 127	19	1.6	621	4	BG497570	BG497570	601859461	c 200	19	1.6	810	2	BE542773	BE542773	601066378	
C 128	19	1.6	622	2	BF183824	BF183824	601842607	c 201	19	1.6	814	5	BQ511498	BQ511498	EST618913	
C 129	19	1.6	622	4	BG504452	BG504452	602552419	c 202	19	1.6	816	7	CK021761	CK021761	AGENCOURT	
C 130	19	1.6	623	4	BG388016	BG388016	602412857	c 203	19	1.6	818	7	CO146594	CO146594	EST821647	
C 131	19	1.6	623	4	BG498368	BG498368	602544159	c 204	19	1.6	820	8	BZ728790	BZ728790	OCEDM35TC	
C 132	19	1.6	624	4	BG505232	BG505232	602551708	c 205	19	1.6	823	4	BG615136	BG615136	602643654	
C 133	19	1.6	625	2	BF211187	BF211187	601812609	c 206	19	1.6	824	8	BE540395	BE540395	601063291	
C 134	19	1.6	631	9	CE450541	CE450541	tigr-g88-	c 207	19	1.6	824	8	BZ123940	BZ123940	CH230-265	
C 135	19	1.6	631	9	CL857851	CL857851	OR_CBA008	c 208	19	1.6	825	5	BU334589	BU334589	603870756	
C 136	19	1.6	632	4	BG283364	BG283364	602407107	c 209	19	1.6	825	9	CC563846	CC563846	CH240-475	
C 137	19	1.6	634	4	BM393058	BM393058	50071-2-4	c 210	19	1.6	833	4	BG495142	BG495142	602540547	
C 138	19	1.6	634	4	BM394598	BM394598	50072-2-4	c 211	19	1.6	835	9	CC985099	CC985099	ZUAAF32TV	
C 139	19	1.6	636	4	BG113556	BG113556	602284026	c 212	19	1.6	840	2	BF697711	BF697711	602131254	
C 140	19	1.6	636	4	BG890072	BG890072	EST515923	c 213	19	1.6	840	8	AQ023855	AQ023855	CPGR0202A	
C 141	19	1.6	638	2	BF210365	BF210365	601874444	c 214	19	1.6	852	4	BG838740	BG838740	G202_G2C0	
C 142	19	1.6	638	9	CPA561574	AJ561574	CrypToSopD	c 215	19	1.6	852	5	BU849461	BU849461	AGENCOURT	
C 143	19	1.6	640	2	BF240947	BF240947	601875528	c 216	19	1.6	854	4	BG395937	BG395937	602458545	
C 144	19	1.6	644	2	BF701490	BF701490	602128906	c 217	19	1.6	860	9	CL684264	CL684264	PR10139A	
C 145	19	1.6	649	8	CC014303	CC014303	PUDDL73TD	c 218	19	1.6	864	2	BF571266	BF571266	602077680	
C 146	19	1.6	655	7	CF368839	CF368839	853555	c 219	19	1.6	865	4	BF979165	BF979165	602147796	
C 147	19	1.6	657	9	CL148078	CL148078	104_327_1	c 220	19	1.6	868	4	BG699472	BG699472	602679167	
C 148	19	1.6	661	9	CE173280	CE173280	tigr-g88-	c 221	19	1.6	868	8	BZ728800	BZ728800	OCEDM35TM	
C 149	19	1.6	663	7	CK005526	CK005526	AGENCOURT	c 222	19	1.6	869	2	BF699654	BF699654	602127065	
C 150	19	1.6	666	2	BE198894	BE198894	u883907.Y	c 223	19	1.6	870	7	CF220506	CF220506	AGENCOURT	
C 151	19	1.6	671	2	BB637812	BB637812	B8637812	c 224	19	1.6	870	7	CO083316	CO083316	GR_Ea48F	
C 152	19	1.6	674	5	BQ783237	BQ783237	UI-R-PFO-	c 225	19	1.6	872	4	BI458621	BI458621	603199738	
C 153	19	1.6	674	8	AZ624238	AZ624238	1M0462113	c 226	19	1.6	881	2	BE727155	BE727155	601563209	
C 154	19	1.6	675	2	AW956821	AW956821	EST368891	c 227	19	1.6	888	4	BG426927	BG426927	602493115	
C 155	19	1.6	677	4	BG167599	BG167599	602345650	c 228	19	1.6	890	4	BI197995	BI197995	602762028	
C 156	19	1.6	677	8	BH838302	BH838302	LMCR10001	c 229	19	1.6	892	6	CA481085	CA481085	AGENCOURT	
C 157	19	1.6	680	5	BU323963	BU323963	603491652	c 230	19	1.6	897	2	BE780431	BE780431	601468978	
C 158	19	1.6	680	9	CC746510	CC746510	ZMMBB012	c 231	19	1.6	897	2	BE782559	BE782559	601465848	
C 159	19	1.6	680	9	CC815867	CC815867	ZMMBB052	c 232	19	1.6	898	2	BF308355	BF308355	601894441	
C 160	19	1.6	686	6	CA351735	CA351735	622865	c 233	19	1.6	898	2	BE894799	BE894799	601436373	
C 161	19	1.6	690	7	CN278845	CN278845	170005328	c 234	19	1.6	899	2	BE781635	BE781635	601467494	
C 162	19	1.6	693	5	EX077130	EX077130	BX077130	c 235	19	1.6	902	9	CL683208	CL683208	PR10136D	
C 163	19	1.6	696	7	CF255442	CF255442	mdvnl28_b	c 236	19	1.6	905	2	BE790261	BE790261	601482484	
C 164	19	1.6	698	7	CR374888	CR374888	CR374888	c 237	19	1.6	907	2	BF125953	BF125953	601762974	
C 165	19	1.6	701	5	BX615248	BX615248	6015248	c 238	19	1.6	916	9	CG863821	CG863821	ZMMBB027	
C 166	19	1.6	704	7	CK349493	CK349493	hgghfa14A	c 239	19	1.6	918	2	BF684662	BF684662	602141026	
C 167	19	1.6	704	7	CK349649	CK349649	hgghfa15H	c 240	19	1.6	924	2	BE891874	BE891874	601435434	
C 168	19	1.6	707	4	BI294878	BI294878	UI-R-DKO-	c 241	19	1.6	938	4	BG330866	BG330866	602431508	
C 169	19	1.6	710	8	AQ327639	AQ327639	nbxb00410	c 242	19	1.6	944	4	BG258470	BG258470	602379949	
C 170	19	1.6	716	4	BF977155	BF977155	602146893	c 243	19	1.6	948	4	BG614975	BG614975	602644216	

C 244	19	1.6	953	2	BE795433	BE795433	601592990	317	18	1.5	356	1	AU278491	AU278491
C 245	19	1.6	957	4	BG330094	BG330094	602428977	318	18	1.5	357	1	AU278456	AU278456
C 246	19	1.6	972	4	BG403393	BG403393	602419069	C 319	18	1.5	357	1	AU278456	AU278456
C 247	19	1.6	978	8	BG390399	BG390399	602416021	C 320	18	1.5	360	1	AU097732	AU097732
C 248	19	1.6	985	8	CC247210	CC247210	CH261-87M	321	18	1.5	360	4	BI537007	BI537007
C 249	19	1.6	1010	7	CN641908	CN641908	ILLUMIN	C 322	18	1.5	365	1	AI895729	AI895729
C 250	19	1.6	1094	8	CC271799	CC271799	CH261-130	C 323	18	1.5	371	1	AV588572	AV588572
C 251	19	1.6	1136	4	BF976431	BF976431	602245238	C 324	18	1.5	371	1	AV588573	AV588573
C 252	19	1.6	1168	8	CC302416	CC302416	CH261-68D	C 325	18	1.5	375	7	CN999388	CN999388
C 253	19	1.6	1288	3	AY222966	AY222966	Schistoso	C 326	18	1.5	376	1	AI593542	AI593542
C 254	19	1.6	1354	4	BG488752	BG488752	602534713	C 327	18	1.5	378	8	CC074529	CC074529
C 255	19	1.6	1626	8	AQ839835	AQ839835	69115-C61	C 328	18	1.5	379	6	BY769384	BY769384
C 256	18	1.5	75	7	CF848676	CF848676	pMA007XL	C 329	18	1.5	379	9	CG743314	CG743314
C 257	18	1.5	122	8	AZ083368	AZ083368	RPCI-23-4	C 330	18	1.5	383	2	AW619733	AW619733
C 258	18	1.5	123	5	BQ255115	BQ255115	NISC-Jp12	C 331	18	1.5	386	5	BP109706	BP109706
C 259	18	1.5	163	8	CR525226	CR525226	CR525226	C 332	18	1.5	386	5	BQ360112	BQ360112
C 260	18	1.5	144	6	CD956530	CD956530	SCD 72 Ge	C 333	18	1.5	386	7	CF615312	CF615312
C 261	18	1.5	149	2	BF876320	BF876320	RC4-ET013	C 334	18	1.5	389	5	BP112766	BP112766
C 262	18	1.5	150	6	BY607486	BY607486	BY607486	C 335	18	1.5	389	6	CD437580	CD437580
C 263	18	1.5	163	2	BS587646	BS587646	BS587646	C 336	18	1.5	390	1	AU234198	AU234198
C 264	18	1.5	163	8	AZ439991	AZ439991	IMO230C19	C 337	18	1.5	390	1	AU234370	AU234370
C 265	18	1.5	186	9	CR078033	CR078033	Reverbe 8	C 338	18	1.5	390	1	AU234416	AU234416
C 266	18	1.5	187	2	AW055817	AW055817	SWAMCAC38	C 339	18	1.5	390	1	AV667673	AV667673
C 267	18	1.5	218	6	CB225057	CB225057	10M28D09	C 340	18	1.5	390	6	CB531256	CB531256
C 268	18	1.5	223	4	BM134711	BM134711	WHE0458D	C 341	18	1.5	390	7	CK958382	CK958382
C 269	18	1.5	224	1	AJ283324	AJ283324	4A3B-AA-	C 342	18	1.5	394	1	AV664689	AV664689
C 270	18	1.5	224	2	BB002559	BB002559	BB002559	C 343	18	1.5	395	6	BY702753	BY702753
C 271	18	1.5	245	7	CV385454	CV385454	QV1-BT063	C 344	18	1.5	396	2	BE072715	BE072715
C 272	18	1.5	246	2	AW033952	AW033952	EST277523	C 345	18	1.5	396	4	BM323882	BM323882
C 273	18	1.5	247	1	AV374437	AV374437	AV374437	C 346	18	1.5	399	7	CK958678	CK958678
C 274	18	1.5	248	7	CN235458	CN235458	WLB083D10	C 347	18	1.5	400	1	AU278599	AU278599
C 275	18	1.5	255	2	BB582899	BB582899	BB582899	C 348	18	1.5	400	1	AU278771	AU278771
C 276	18	1.5	264	7	CV344233	CV344233	MR1-DT005	C 349	18	1.5	400	1	AU278817	AU278817
C 277	18	1.5	265	1	AV238273	AV238273	AV238273	C 350	18	1.5	403	6	CB812459	CB812459
C 278	18	1.5	273	1	AI473225	AI473225	th57b03.x	C 351	18	1.5	403	7	CK960485	CK960485
C 279	18	1.5	273	5	BP155520	BP155520	BP155520	C 352	18	1.5	405	7	CO000987	CO000987
C 280	18	1.5	274	7	N892939	N892939	K89293F Huma	C 353	18	1.5	407	1	AV592320	AV592320
C 281	18	1.5	275	2	BE154374	BE154374	FM2-HT034	C 354	18	1.5	407	4	BG939375	BG939375
C 282	18	1.5	276	7	CO941583	CO941583	UMC-peov3	C 355	18	1.5	407	6	CB771441	CB771441
C 283	18	1.5	286	1	AI145162	AI145162	UI-R-BT0-	C 356	18	1.5	409	1	AV662643	AV662643
C 284	18	1.5	288	1	AA007210	AA007210	13CDNA50B	C 357	18	1.5	410	4	BM435471	BM435471
C 285	18	1.5	288	2	BB074782	BB074782	BB074782	C 358	18	1.5	411	7	H02448	H02448
C 286	18	1.5	288	7	CO508900	CO508900	tah28H03.	C 359	18	1.5	414	4	BM436190	BM436190
C 287	18	1.5	291	7	CK095043	CK095043	1066P49.3	C 360	18	1.5	414	5	BP100662	BP100662
C 288	18	1.5	294	8	CC102721	CC102721	MR3-K34.1	C 361	18	1.5	414	7	CO520679	CO520679
C 289	18	1.5	295	4	BI054066	BI054066	CR3-GN034	C 362	18	1.5	415	1	AV663976	AV663976
C 290	18	1.5	300	1	AJ462076	AJ462076	AJ462076	C 363	18	1.5	417	1	AV597070	AV597070
C 291	18	1.5	302	2	BB470743	BB470743	BB470743	C 364	18	1.5	417	2	BE322297	BE322297
C 292	18	1.5	302	6	CB700687	CB700687	AMGNNUC:H	C 365	18	1.5	418	1	AV814373	AV814373
C 293	18	1.5	303	2	BE631030	BE631030	uu52d09.x	C 366	18	1.5	418	4	BI922788	BI922788
C 294	18	1.5	303	8	AZ005096	AZ005096	RPCI-23-3	C 367	18	1.5	418	9	CE774669	CE774669
C 295	18	1.5	305	6	CD605291	CD605291	RK020A3H0	C 368	18	1.5	419	1	AJ722806	AJ722806
C 296	18	1.5	318	9	CL224893	CL224893	ZMMBC054	C 369	18	1.5	419	1	AV665841	AV665841
C 297	18	1.5	319	1	AV664982	AV664982	AV664982	C 370	18	1.5	419	1	AA315557	AA315557
C 298	18	1.5	323	7	CK729792	CK729792	UMC-bdv_0	C 371	18	1.5	420	1	AU232623	AU232623
C 299	18	1.5	331	7	CN654088	CN654088	UMC-bcl_0	C 372	18	1.5	420	7	R00194	R00194
C 300	18	1.5	332	7	CO781601	CO781601	BL012D_E1	C 373	18	1.5	421	1	AV663337	AV663337
C 301	18	1.5	336	2	BB553964	BB553964	BB553964	C 374	18	1.5	421	1	AV664366	AV664366
C 302	18	1.5	336	6	CB226935	CB226935	1Ru32C02	C 375	18	1.5	421	1	AV664489	AV664489
C 303	18	1.5	337	7	CF116007	CF116007	et3-42.z1	C 376	18	1.5	422	5	BQ666647	BQ666647
C 304	18	1.5	347	1	AA281917	AA281917	zt07f05.x	C 377	18	1.5	423	2	BF430378	BF430378
C 305	18	1.5	347	5	BP109488	BP109488	BP109488	C 378	18	1.5	423	5	BQ666719	BQ666719
C 306	18	1.5	347	7	CO94271	CO94271	UMC-pd12	C 379	18	1.5	426	2	AW297996	AW297996
C 307	18	1.5	349	6	CB225957	CB225957	1RT25F03-	C 380	18	1.5	427	1	AV597363	AV597363
C 308	18	1.5	350	1	AU231430	AU231430	AU231430	C 381	18	1.5	427	6	CA398900	CA398900
C 309	18	1.5	350	1	AU233883	AU233883	AU233883	C 382	18	1.5	428	2	BF430325	BF430325
C 310	18	1.5	350	1	AU234055	AU234055	AU234055	C 383	18	1.5	431	1	AV615111	AV615111
C 311	18	1.5	350	1	AU234090	AU234090	AU234090	C 384	18	1.5	432	6	CB435375	CB435375
C 312	18	1.5	351	1	AV615767	AV615767	AV615767	C 385	18	1.5	433	1	AI894418	AI894418
C 313	18	1.5	351	7	CR474305	CR474305	CR474305	C 386	18	1.5	433	2	AW524835	AW524835
C 314	18	1.5	352	1	AI559177	AI559177	tg42f03.x	C 387	18	1.5	435	1	AV588885	AV588885
C 315	18	1.5	352	2	BP652938	BP652938	276657 MA	C 388	18	1.5	435	1	AV593075	AV593075
C 316	18	1.5	355	1	AU278284	AU278284	AU278284	C 389	18	1.5	435	4	BG933446	BG933446

C 390	18	1.5	435	5	BP106205	BP106205	463	18	1.5	509	1	AV607463	AV607463
C 391	18	1.5	435	6	CA006413	CA006413	464	18	1.5	509	1	AV609810	AV609810
C 392	18	1.5	435	6	CD287287	15_J13_ab	C 465	18	1.5	511	1	AJ670992	AJ670992
C 393	18	1.5	437	1	AA809058	nw17a04.s	C 466	18	1.5	511	4	BM324471	BM324471
C 394	18	1.5	437	6	CB537603	775251_MA	C 467	18	1.5	511	7	CO309709	CO309709
C 395	18	1.5	437	7	CF930246	CF-03-R-D	C 468	18	1.5	512	6	CD443012	CD443012
C 396	18	1.5	438	2	BE601076	PF1_96_F0	C 469	18	1.5	514	2	BE367133	BE367133
C 397	18	1.5	442	5	BX565821	BX565821	C 470	18	1.5	515	6	CB046799	CB046799
C 398	18	1.5	446	1	AV616390	AV616390	C 471	18	1.5	516	6	CD444523	CD444523
C 399	18	1.5	446	7	CF930628	CF-04-R-	C 472	18	1.5	516	4	BI400921	BI400921
C 400	18	1.5	447	1	AV590167	AV590167	C 473	18	1.5	516	4	BJ029616	BJ029616
C 401	18	1.5	448	1	AV597055	AV597055	C 474	18	1.5	516	6	CA672290	CA672290
C 402	18	1.5	449	5	BP100595	BP100595	C 475	18	1.5	517	8	AZ468953	AZ468953
C 403	18	1.5	451	5	BP106319	BP106319	C 476	18	1.5	517	8	AZ478756	AZ478756
C 404	18	1.5	452	4	BM431938	1JB14D1.	C 477	18	1.5	518	6	CB170925	CB170925
C 405	18	1.5	454	7	CO892719	BovGer_21	C 478	18	1.5	518	6	CD446119	CD446119
C 406	18	1.5	455	6	CD433510	CD433510	C 479	18	1.5	519	9	CE206076	CE206076
C 407	18	1.5	456	1	AJ696982	AJ696982	C 480	18	1.5	520	2	BE754622	BE754622
C 408	18	1.5	457	1	AV617419	AV617419	C 481	18	1.5	520	4	BM130246	BM130246
C 409	18	1.5	457	2	AW129868	707003G06	C 482	18	1.5	520	6	CB532045	CB532045
C 410	18	1.5	457	4	BM430594	1Du001G05	C 483	18	1.5	521	5	BO584448	BO584448
C 411	18	1.5	457	6	CB538027	CB538027	C 484	18	1.5	521	7	CK961666	CK961666
C 412	18	1.5	458	6	CB464366	725571_MA	C 485	18	1.5	523	2	BB764381	BB764381
C 413	18	1.5	459	6	CB535499	768931_MA	C 486	18	1.5	523	4	BM130161	BM130161
C 414	18	1.5	460	7	CO531042	CO531042	C 487	18	1.5	523	8	AZ004941	AZ004941
C 415	18	1.5	460	8	CO931814	CSU-K34.1	C 488	18	1.5	524	5	BQ168626	BQ168626
C 416	18	1.5	461	7	CF923251	CF-01-R-C	C 489	18	1.5	525	1	AU279041	AU279041
C 417	18	1.5	462	6	CB536292	770644_MA	C 490	18	1.5	525	6	CD444531	CD444531
C 418	18	1.5	463	2	BB370203	BB370203	C 491	18	1.5	526	5	BP105653	BP105653
C 419	18	1.5	467	1	AJ691083	AJ691083	C 492	18	1.5	526	6	CD443064	CD443064
C 420	18	1.5	467	1	AJ696782	AJ696782	C 493	18	1.5	526	9	CL412159	CL412159
C 421	18	1.5	467	4	BM433771	10M04007a	C 494	18	1.5	527	6	CD765429	CD765429
C 422	18	1.5	468	8	BH052665	RPCI-24-3	C 495	18	1.5	527	7	CK574274	CK574274
C 423	18	1.5	471	2	BP429683	BP429683	C 496	18	1.5	527	7	CK982038	CK982038
C 424	18	1.5	471	6	CD203592	LG_AML_10	C 497	18	1.5	528	9	CL322039	CL322039
C 425	18	1.5	472	6	CB535714	769947_MA	C 498	18	1.5	529	4	BI295592	BI295592
C 426	18	1.5	473	2	BE367575	PF1_9_F06	C 499	18	1.5	532	5	BP109468	BP109468
C 427	18	1.5	473	6	CB222654	11L3J0D05	C 500	18	1.5	532	6	CB535210	CB535210
C 428	18	1.5	475	5	BU039862	PP_L8a000	C 501	18	1.5	533	4	BJ040404	BJ040404
C 429	18	1.5	475	5	BU686154	UI-CF-DUL	C 502	18	1.5	533	5	BP102823	BP102823
C 430	18	1.5	479	2	BE367479	PF1_8_F02	C 503	18	1.5	534	6	CB536002	CB536002
C 431	18	1.5	479	2	BE451375	EST402263	C 504	18	1.5	534	6	CB536140	CB536140
C 432	18	1.5	480	2	BE597419	PF1_69_C1	C 505	18	1.5	535	2	AW629326	AW629326
C 433	18	1.5	481	1	AI058429	UI-R-CI-k	C 506	18	1.5	535	2	BE363981	BE363981
C 434	18	1.5	481	9	CC860600	NDL_131G7	C 507	18	1.5	536	9	CR186770	CR186770
C 435	18	1.5	482	6	CD440238	EL01N0552	C 508	18	1.5	537	6	CB717642	CB717642
C 436	18	1.5	482	8	BZ884750	CH240_189	C 509	18	1.5	537	6	CD203058	CD203058
C 437	18	1.5	483	6	CB536249	770600_MA	C 510	18	1.5	538	4	BI301230	BI301230
C 438	18	1.5	483	7	CF930539	CF--04-R-	C 511	18	1.5	539	2	AW678619	AW678619
C 439	18	1.5	484	5	BP101837	BP101837	C 512	18	1.5	540	6	CB531456	CB531456
C 440	18	1.5	484	7	CF931053	CF--06-R-	C 513	18	1.5	541	6	CA005815	CA005815
C 441	18	1.5	485	9	CL871127	abe74d06.	C 514	18	1.5	544	5	BU999927	BU999927
C 442	18	1.5	487	7	CF634742	znrrw000.0	C 515	18	1.5	544	8	AQ665442	AQ665442
C 443	18	1.5	487	7	CR383067	CR383067	C 516	18	1.5	545	2	AW746936	AW746936
C 444	18	1.5	488	2	BF770251	RC1-IT001	C 517	18	1.5	546	6	CD057057	CD057057
C 445	18	1.5	488	5	BP100548	BP100548	C 518	18	1.5	546	6	CD443309	CD443309
C 446	18	1.5	490	8	BH272241	CH230-162	C 519	18	1.5	547	1	AI3033310	AI3033310
C 447	18	1.5	494	2	BF600611	265346_MA	C 520	18	1.5	548	4	BM325875	BM325875
C 448	18	1.5	494	7	CR383550	CR383550	C 521	18	1.5	548	4	BM325875	BM325875
C 449	18	1.5	495	6	CD442047	CD442047	C 522	18	1.5	548	7	CO885570	CO885570
C 450	18	1.5	496	2	BF653234	276936_MA	C 523	18	1.5	549	9	CE250123	CE250123
C 451	18	1.5	497	1	AV612820	AV612820	C 524	18	1.5	550	2	AW662644	AW662644
C 452	18	1.5	497	6	CD442580	EL01N0413	C 525	18	1.5	550	1	AW671887	AW671887
C 453	18	1.5	498	2	BF599523	263339_MA	C 526	18	1.5	550	7	BP106128	BP106128
C 454	18	1.5	500	2	BE367652	PF1_9_D04	C 527	18	1.5	550	7	CF615317	CF615317
C 455	18	1.5	500	6	CD605391	CD605391	C 528	18	1.5	553	1	AU278940	AU278940
C 456	18	1.5	500	8	AZ020126	RPCI-23-3	C 529	18	1.5	554	2	BF727720	BF727720
C 457	18	1.5	502	1	AV611923	AV611923	C 530	18	1.5	555	6	CD426863	CD426863
C 458	18	1.5	503	6	CD286136	10_L1_abd	C 531	18	1.5	558	7	CN434013	CN434013
C 459	18	1.5	503	5	BP046868	BP046868	C 532	18	1.5	558	7	CF244498	CF244498
C 460	18	1.5	504	4	BM434203	BM434203	C 533	18	1.5	559	7	CF244498	CF244498
C 461	18	1.5	507	2	BF005991	EST434489	C 534	18	1.5	563	4	BM322997	BM322997
C 462	18	1.5	507	2	BF005991	EST434489	C 535	18	1.5	563	8	AQ456627	AQ456627

C 536	18	1.5	563	9	CE093704	CR093704	tigr-ges-	609	18	1.5	604	7	CK967566	CK967566	4083113 B
537	18	1.5	564	2	BE367614	PII_9_B10		610	18	1.5	605	4	BM607869	BM607869	170006870
538	18	1.5	564	4	BI880188	fm73h02.x		611	18	1.5	605	7	CO783442	CO783442	BL018A A1
539	18	1.5	564	5	BU451322	403771519		612	18	1.5	606	5	BP104707	BP104707	BP104707
540	18	1.5	566	7	CK958509	4099243 B		613	18	1.5	606	6	CB535834	CB535834	770080 MA
541	18	1.5	567	8	AZ245932	RPCI-23-1		614	18	1.5	606	9	CE144005	CE144005	tigr-ges-
542	18	1.5	568	4	BI922962	ES7542866		615	18	1.5	607	6	CB538107	CB538107	778068 MA
543	18	1.5	568	6	CD925153	G750.116A		616	18	1.5	607	6	CD214855	CD214855	pgmnp.pko
544	18	1.5	569	6	CB466591	732133 MA		617	18	1.5	608	6	CB446376	CB446376	700079 MA
545	18	1.5	569	6	CD443346	EL01N0425		618	18	1.5	608	6	CB534823	CB534823	768183 MA
546	18	1.5	571	2	BF667925	602122164		619	18	1.5	608	6	CB535535	CB535535	768974 MA
547	18	1.5	571	7	CK946434	4070723 B		620	18	1.5	608	6	CB537984	CB537984	775938 MA
548	18	1.5	572	5	BN967941	LM24HW008		621	18	1.5	608	8	AZ956059	AZ956059	2M0236G23
549	18	1.5	572	6	CB536859	771923 MA		622	18	1.5	609	6	CB433841	CB433841	610278 MA
550	18	1.5	573	8	BH204048	Sml-48C9		623	18	1.5	609	7	CF764953	CF764953	CS5003102
551	18	1.5	575	4	BM326394	PICI_56 E		624	18	1.5	609	7	CO533213	CO533213	3530_1_21
552	18	1.5	576	1	AV617991	AV617991		625	18	1.5	610	7	CK974232	CK974232	4105054 B
553	18	1.5	577	8	AZ857337	2M0162B08		626	18	1.5	610	7	CO886392	CO886392	BoyGen_14
554	18	1.5	578	4	BM429478	LA22H04 B		627	18	1.5	611	2	BE367671	BE367671	PII_9_F06
555	18	1.5	578	8	AZ118333	RPCI-23-4		628	18	1.5	611	6	CB447887	CB447887	701926 MA
556	18	1.5	578	8	BH235837	ATZKF60TF		629	18	1.5	611	6	CB534903	CB534903	768275 MA
557	18	1.5	579	2	BF429617	11974 MAR		630	18	1.5	611	7	CK962923	CK962923	4077395 B
558	18	1.5	579	6	CD734732	4049011 1		631	18	1.5	614	5	BP101258	BP101258	BP101258
559	18	1.5	579	9	CC565357	CH240_478		632	18	1.5	614	6	CB534368	CB534368	767693 MA
560	18	1.5	580	8	BH100267	RPCI-24-3		633	18	1.5	615	6	CB452394	CB452394	707256 MA
561	18	1.5	581	2	BE568891	601342259		634	18	1.5	615	6	CB461850	CB461850	721866 MA
562	18	1.5	582	5	BP109442	BP109442		635	18	1.5	615	6	CB464745	CB464745	725975 MA
563	18	1.5	582	5	BP327195	BP327195		636	18	1.5	615	9	CR122221	CR122221	Forward s
564	18	1.5	582	7	CO523264	3530_1_15		637	18	1.5	615	9	CE839105	CE839105	tigr-ges-
565	18	1.5	583	6	CD433549	EL01N0427		638	18	1.5	615	6	CB429979	CB429979	605833 MA
566	18	1.5	584	2	BF429651	1438 MARC		639	18	1.5	616	6	CB446015	CB446015	697798 MA
567	18	1.5	584	2	CO531041	3530_1_20		640	18	1.5	616	6	CB447534	CB447534	701542 MA
568	18	1.5	584	7	CR553699	CR553699		641	18	1.5	616	6	CB530882	CB530882	742113 MA
569	18	1.5	584	7	CL760363	OR_BBA012		642	18	1.5	616	6	CB533520	CB533520	761506 MA
570	18	1.5	585	4	BI922970	EST542874		643	18	1.5	617	6	CB433187	CB433187	609510 MA
571	18	1.5	585	5	BP104558	BP104558		644	18	1.5	617	6	CB534484	CB534484	767819 MA
572	18	1.5	585	7	CF244497	3530_1_22		645	18	1.5	617	7	CN788259	CN788259	4122572 B
573	18	1.5	585	9	CR023267	Forward s		646	18	1.5	617	8	AQ279502	AQ279502	CTIBI-EI-
574	18	1.5	586	5	BN956569	LM24HW006		647	18	1.5	618	6	CA733537	CA733537	wplc.pko
575	18	1.5	586	7	CF766908	CE5002436		648	18	1.5	618	6	CB461840	CB461840	721854 MA
576	18	1.5	587	7	CF613650	CE5007932		649	18	1.5	618	7	CN651807	CN651807	EG_CWGRS
577	18	1.5	588	2	BE597194	PII_69_C1		650	18	1.5	618	8	CR822431	CR822431	Oa_sp1bn
578	18	1.5	590	6	CD433933	EL01N0316		651	18	1.5	618	8	AZ566730	AZ566730	225Pv801
579	18	1.5	591	7	CK979517	4111001 B		652	18	1.5	619	5	BP112811	BP112811	BE030013A
580	18	1.5	591	9	CL579393	OB_BA003		653	18	1.5	619	7	CK940185	CK940185	4113496 B
581	18	1.5	592	7	CF765290	CE5002198		654	18	1.5	620	6	CB452259	CB452259	707113 MA
582	18	1.5	593	2	AM678759	WS1_1_B04		655	18	1.5	620	6	CB533818	CB533818	765267 MA
583	18	1.5	593	6	CB468497	734337 MA		656	18	1.5	621	4	BG690060	BG690060	733486 MA
584	18	1.5	593	7	CF765295	CE5002208		657	18	1.5	621	6	CB537797	CB537797	775453 MA
585	18	1.5	593	8	AZ904496	RPCI-24-1		658	18	1.5	621	8	BH097162	BH097162	RPCI-24-2
586	18	1.5	594	6	CB464189	725381 MA		659	18	1.5	622	7	CA435123	CA435123	BE030013A
587	18	1.5	594	6	CD436248	EL01N0372		660	18	1.5	622	2	BF006564	BF006564	EST435062
588	18	1.5	594	6	CD436771	EL01N0363		661	18	1.5	623	6	CB467724	CB467724	733486 MA
589	18	1.5	595	5	BO667404	p661804_Y		662	18	1.5	623	6	CD422950	CD422950	SAL_38_F0
590	18	1.5	595	5	BH248824	603592487		663	18	1.5	623	6	CD444226	CD444226	EL01N0437
591	18	1.5	595	6	CA211092	SCPSB113		664	18	1.5	624	5	BQ389849	BQ389849	NISC_mq09
592	18	1.5	596	8	AZ380603	IM0136B10		665	18	1.5	624	6	CA401218	CA401218	EL01N0416
593	18	1.5	596	7	KN034698	Math_P2_J		666	18	1.5	624	6	CB463298	CB463298	723671 MA
594	18	1.5	596	7	KN854053	Ha_mx0_20		667	18	1.5	624	6	CB535455	CB535455	768879 MA
595	18	1.5	598	4	BI696127	603345962		668	18	1.5	624	9	DR22P13T	DR22P13T	Danio rer
596	18	1.5	599	6	CB582670	AMGNUC:N		669	18	1.5	625	7	CF244891	CF244891	3530_1_5
597	18	1.5	599	7	CV093544	FAMU_USDA		670	18	1.5	625	9	DR140118	DR140118	Danio rer
598	18	1.5	600	2	BF107897	601823917		671	18	1.5	626	6	CB531472	CB531472	743993 MA
599	18	1.5	600	6	CA630943	wleln.pko		672	18	1.5	626	7	CF763913	CF763913	CE5004964
600	18	1.5	600	6	CB538113	776074 MA		673	18	1.5	626	7	CF764404	CF764404	CE5003975
601	18	1.5	600	7	CF762283	CE5000529		674	18	1.5	629	6	CD445210	CD445210	EL01N0448
602	18	1.5	602	6	CB425654	600736 MA		675	18	1.5	630	6	CB515829	CB515829	sealr9b50
603	18	1.5	602	6	CB537823	775481 MA		676	18	1.5	630	7	CK976211	CK976211	4107650 B
604	18	1.5	602	6	CD435268	EL01N0357		677	18	1.5	630	7	AG103982	AG103982	Pan_trog1
605	18	1.5	602	6	CD443016	EL01N0421		678	18	1.5	630	9	AG103982	AG103982	Pan_trog1
606	18	1.5	604	7	CF763021	CE5006750		679	18	1.5	631	7	CK962186	CK962186	4076728 B
607	18	1.5	604	7	CK960421	4101393 B		680	18	1.5	633	6	CD444719	CD444719	EL01N0443
608	18	1.5	604	7	CK962182	4076726 B		681	18	1.5	633	7	CN655096	CN655096	SAL_US005

682	18	1.5	634	2	BE600449	BB600449	PIL_96_F0	C 755	18	1.5	673	9	CL360124	CL360124	RPCI44_33
683	18	1.5	634	7	CF767201	CF767201	CES002334	C 756	18	1.5	674	6	CD444169	CD444169	EL01N0436
684	18	1.5	634	7	CK569673	CK569673	HO14A15W	C 757	18	1.5	675	6	CD433166	CD433166	EL01N0305
685	18	1.5	634	7	CO517963	CO517963	3530_1_11	C 758	18	1.5	675	6	CD435749	CD435749	EL01N0365
686	18	1.5	636	8	AZ300559	AZ300559	RPCI_23-1	C 759	18	1.5	675	6	CD436286	CD436286	EL01N0372
687	18	1.5	636	9	CE540830	CE540830	tigr-g88-	C 760	18	1.5	675	6	CF068679	CF068679	EST669400
688	18	1.5	636	9	CE545436	CE545436	tigr-g88-	C 761	18	1.5	675	7	CF763619	CF763619	CES005568
689	18	1.5	637	2	BE367243	BE367243	PIL_43_D0	C 762	18	1.5	675	8	BH330562	BH330562	AG-ND-130
690	18	1.5	637	6	CB532731	CB532731	757210_B	C 763	18	1.5	676	6	CD051968	CD051968	PLX203_Ca
691	18	1.5	637	7	CK972648	CK972648	4103307_B	C 764	18	1.5	677	1	AI207462	AI207462	HA2813_Hu
692	18	1.5	638	6	CD435954	CD435954	EL01N0368	C 765	18	1.5	677	6	CB533795	CB533795	765243_MA
693	18	1.5	638	9	CE149312	CE149312	tigr-g88-	C 766	18	1.5	677	6	CD448077	CD448077	EL01N0204
694	18	1.5	639	6	CD442177	CD442177	EL01N0406	C 767	18	1.5	677	9	CE432280	CE432280	tigr-g88-
695	18	1.5	641	7	CO872722	CO872722	BoyGen_01	C 768	18	1.5	678	6	CB533693	CB533693	765120_MA
696	18	1.5	642	7	CF769007	CF769007	CES001595	C 769	18	1.5	678	6	CD444249	CD444249	EL01N0437
697	18	1.5	642	7	CK967223	CK967223	4082716_B	C 770	18	1.5	680	6	CD444231	CD444231	EL01N0437
698	18	1.5	643	4	BJ467321	BJ467321	BJ467321	C 771	18	1.5	680	6	CD445222	CD445222	EL01N0449
699	18	1.5	644	6	CB467765	CB467765	733528_MA	C 772	18	1.5	680	7	CNS53729	CNS53729	UI-M-HOO-
700	18	1.5	644	7	CF767395	CF767395	CES000239	C 773	18	1.5	680	7	CO533119	CO533119	3530_1_21
701	18	1.5	644	7	CF943974	CF943974	TtEST-A02	C 774	18	1.5	681	7	CF625956	CF625956	zmrw805_0
702	18	1.5	644	9	CE313186	CE313186	tigr-g88-	C 775	18	1.5	682	9	CB051523	CB051523	tigr-g88-
703	18	1.5	645	7	CK949139	CK949139	4073974_B	C 776	18	1.5	683	2	BB538928	BB538928	BBS38928
704	18	1.5	646	6	CB533519	CB533519	761505_MA	C 777	18	1.5	683	6	CD442370	CD442370	EL01N0408
705	18	1.5	646	6	CD445010	CD445010	EL01N0446	C 778	18	1.5	683	7	CN653624	CN653624	Eg_CWGRS
706	18	1.5	646	7	CF762944	CF762944	CES006990	C 779	18	1.5	684	7	CF627176	CF627176	zmrw805_0
707	18	1.5	647	6	CD442823	CD442823	EL01N0418	C 780	18	1.5	684	7	CF925233	CF925233	ML MQ1_05
708	18	1.5	649	6	CB453604	CB453604	709483_MA	C 781	18	1.5	685	6	CB533375	CB533375	757992_MA
709	18	1.5	649	7	CF728498	CF728498	UI-M-HCO-	C 782	18	1.5	685	7	CN651880	CN651880	Eg_CWGRS
710	18	1.5	649	7	CF764211	CF764211	CES004377	C 783	18	1.5	686	6	CB457756	CB457756	715868_MA
711	18	1.5	649	7	CN823018	CN823018	Oa_splbn	C 784	18	1.5	686	7	CF766703	CF766703	CES007380
712	18	1.5	650	6	CD444871	CD444871	EL01N0445	C 785	18	1.5	686	7	CK960672	CK960672	411603_B
713	18	1.5	650	6	CE778595	CE778595	tigr-g88-	C 786	18	1.5	686	9	CL445861	CL445861	ZMMBB046
714	18	1.5	651	6	CD365558	CD365558	EL01N0358	C 787	18	1.5	687	2	AW216613	AW216613	EST295327
715	18	1.5	651	7	CN651633	CN651633	Eg_CWGRS	C 788	18	1.5	687	2	AW651151	AW651151	EST373223
716	18	1.5	651	8	BZ231502	BZ231502	CH230-438	C 789	18	1.5	688	7	CK982425	CK982425	4114699_B
717	18	1.5	652	8	AZ567303	AZ567303	231PvH01	C 790	18	1.5	690	6	CD446842	CD446842	EL01T0207
718	18	1.5	654	6	CD343039	CD343039	EL01N0318	C 791	18	1.5	690	7	CO519377	CO519377	3530_1_12
719	18	1.5	654	7	CF767830	CF767830	CES005863	C 792	18	1.5	690	9	CL547766	CL547766	OB_Ba008
720	18	1.5	654	7	CN653386	CN653386	Eg_CWGRS	C 793	18	1.5	691	1	AV920342	AV920342	AV920342
721	18	1.5	655	9	CE148653	CE148653	tigr-g88-	C 794	18	1.5	691	6	CD443900	CD443900	EL01N0432
722	18	1.5	655	2	AW621113	AW621113	707003G06	C 795	18	1.5	691	6	CD444857	CD444857	EL01N0444
723	18	1.5	656	5	BP107231	BP107231	BP107231	C 796	18	1.5	691	7	CN604369	CN604369	USDA_FP_1
724	18	1.5	656	6	CD363632	CD363632	EL01N0369	C 797	18	1.5	691	7	CN653707	CN653707	Eg_CWGRS
725	18	1.5	657	7	CN823073	CN823073	Oa_splbn	C 798	18	1.5	692	6	CB460522	CB460522	720414_MA
726	18	1.5	657	8	AQ268874	AQ268874	RPCI11-77	C 799	18	1.5	692	6	CB531798	CB531798	754665_MA
727	18	1.5	658	6	CD466000	CD466000	EL01T0206	C 800	18	1.5	692	6	CB531909	CB531909	754781_MA
728	18	1.5	658	7	CO887052	CO887052	BoyGen_15	C 801	18	1.5	692	9	CE822637	CE822637	tigr-g88-
729	18	1.5	659	7	CF764057	CF764057	CES004858	C 802	18	1.5	693	6	CB531099	CB531099	742349_MA
730	18	1.5	659	7	CF767798	CF767798	CES006110	C 803	18	1.5	693	6	CD443993	CD443993	EL01N0433
731	18	1.5	661	6	CD435329	CD435329	EL01N0358	C 804	18	1.5	693	7	CN820532	CN820532	HRO4413_D
732	18	1.5	661	7	CF629293	CF629293	zmrw848_0	C 805	18	1.5	693	9	CE406370	CE406370	tigr-g88-
733	18	1.5	661	9	AG106639	AG106639	Pan_trog1	C 806	18	1.5	694	7	CF114305	CF114305	ShultZomi
734	18	1.5	662	1	AL632156	AL632156	AL632156	C 807	18	1.5	694	7	CK981743	CK981743	4113880_B
735	18	1.5	663	6	CD444968	CD444968	EL01N0446	C 808	18	1.5	694	7	CN653459	CN653459	Eg_CWGRS
736	18	1.5	664	6	CB533084	CB533084	757658_MA	C 809	18	1.5	695	6	CB533396	CB533396	758016_MA
737	18	1.5	664	6	CD444644	CD444644	EL01N0442	C 810	18	1.5	695	6	CB533694	CB533694	765121_MA
738	18	1.5	664	7	CO892220	CO892220	BoyGen_20	C 811	18	1.5	695	7	CK965876	CK965876	4081190_B
739	18	1.5	664	9	AG120126	AG120126	Pan_trog1	C 812	18	1.5	695	7	CO520680	CO520680	3530_1_13
740	18	1.5	665	6	CB465694	CB465694	727757_MA	C 813	18	1.5	695	7	CO528497	CO528497	3530_1_18
741	18	1.5	665	7	CF765616	CF765616	CES001661	C 814	18	1.5	695	7	CR370059	CR370059	CR370059
742	18	1.5	665	7	CO533214	CO533214	3530_1_21	C 815	18	1.5	695	9	CL195333	CL195333	104_420_1
743	18	1.5	666	1	AV921560	AV921560	AV921560	C 816	18	1.5	696	7	CK836968	CK836968	4061978_B
744	18	1.5	666	4	BJ261755	BJ261755	BJ261755	C 817	18	1.5	696	7	CK967960	CK967960	4083497_B
745	18	1.5	666	5	BU381206	BU381206	603582211	C 818	18	1.5	697	7	CK150947	CK150947	EST780262
746	18	1.5	666	7	CK958766	CK958766	4099627_B	C 819	18	1.5	698	4	BG440255	BG440255	GA_Ea000
747	18	1.5	667	6	BY735093	BY735093	BY735093	C 820	18	1.5	698	7	CK983251	CK983251	4076554_B
748	18	1.5	667	6	CB464645	CB464645	725867_MA	C 821	18	1.5	699	7	CK962543	CK962543	4077112_B
749	18	1.5	667	7	CN653583	CN653583	Eg_CWGRS	C 822	18	1.5	699	7	CK979096	CK979096	4110617_B
750	18	1.5	668	2	AW288747	AW288747	707003G06	C 823	18	1.5	699	7	CO877430	CO877430	BoyGen_05
751	18	1.5	670	7	CN432877	CN432877	B5030003B	C 824	18	1.5	700	7	CN653653	CN653653	Eg_CWGRS
752	18	1.5	671	6	CD443079	CD443079	EL01N0421	C 825	18	1.5	700	7	CN822760	CN822760	Oa_splbn-
753	18	1.5	673	3	CD444339	CD444339	EL01N0438	C 826	18	1.5	701	5	BP112534	BP112534	BF112534
754	18	1.5	673	9	CC472639	CC472639	CH240_296	C 827	18	1.5	703	9	CE272509	CE272509	tigr-g88-

C 828	18	1.5	704	5	BW455207	BW455207	901	18	1.5	760	7	CO522485	CO522485	3530_1_14
C 829	18	1.5	704	6	CD434510	EL01N0324	C 902	18	1.5	762	6	CD437853	CD437853	EL01N0505
C 830	18	1.5	705	6	CA079108	SCSBAM108	C 903	18	1.5	762	6	CN823856	Oa_sp1bn	CN823856
C 831	18	1.5	705	7	CN787127	4121209 B	C 904	18	1.5	764	2	BE307740	601098017	BE307740
C 832	18	1.5	706	6	CB460534	720426 MA	C 905	18	1.5	764	2	CO522486	3530_1_14	CO522486
C 833	18	1.5	706	6	CB468630	734477 MA	C 906	18	1.5	765	6	CD447907	EL01N0202	CD447907
C 834	18	1.5	706	7	CK962541	4077110 B	C 907	18	1.5	767	6	CA403242	EL01N0448	CA403242
C 835	18	1.5	706	7	CO886849	BovGen_15	C 908	18	1.5	767	9	CC919892	CO28911ba	CC919892
C 836	18	1.5	707	4	BU037422	BU037422	C 909	18	1.5	768	7	CB170632	QVK602600	CB170632
C 837	18	1.5	707	7	CK967418	4082457 B	C 910	18	1.5	768	7	CN824128	Oa_sp1bn	CN824128
C 838	18	1.5	707	8	BZ284453	CH230-385	C 911	18	1.5	769	7	CO528496	3530_1_18	CO528496
C 839	18	1.5	707	9	CG310630	CG2CN33TV	C 912	18	1.5	769	8	BZ435744	BONBG30TF	BZ435744
C 840	18	1.5	708	5	BW431154	BW431154	C 913	18	1.5	771	6	CD440418	EL01N0554	CD440418
C 841	18	1.5	708	6	CB972100	CAB10006	C 914	18	1.5	771	9	CR217774	Reverse s	CR217774
C 842	18	1.5	709	6	CB463199	723563 MA	C 915	18	1.5	772	6	CD434074	EL01N0318	CD434074
C 843	18	1.5	711	6	CK967010	4082332 B	C 916	18	1.5	773	8	AQ501223	226B11 mt	AQ501223
C 844	18	1.5	712	6	CD437498	EL01N0501	C 917	18	1.5	774	7	CO533118	3530_1_21	CO533118
C 845	18	1.5	712	7	CK972342	4102923 B	C 918	18	1.5	774	8	BZ163721	CH230-258	BZ163721
C 846	18	1.5	713	6	CB531813	CK972342	C 919	18	1.5	776	6	CA339110	EL01N0314	CA339110
C 847	18	1.5	713	7	CO206734	CO206734	C 920	18	1.5	776	6	CD441214	EL01N0553	CD441214
C 848	18	1.5	713	9	CR795998	GROAAA13D	C 921	18	1.5	777	7	CN823269	Oa_sp1bn	CN823269
C 849	18	1.5	714	6	CD435293	EL01N0358	C 922	18	1.5	778	6	CD434432	EL01N0324	CD434432
C 850	18	1.5	714	6	CD448084	EL01N0204	C 923	18	1.5	778	7	CN824344	Oa_sp1bn	CN824344
C 851	18	1.5	715	4	BM582040	170006872	C 924	18	1.5	778	7	CR374328	CR374328	CR374328
C 852	18	1.5	717	2	AW333556	S2C3 AGS	C 925	18	1.5	780	7	CO646530	ILLUMIGEN	CO646530
C 853	18	1.5	718	6	CD436202	EL01N0372	C 926	18	1.5	781	2	BF266586	HV_CBA001	BF266586
C 854	18	1.5	719	5	BW114369	BW114369	C 927	18	1.5	782	6	CD438353	EL01N0512	CD438353
C 855	18	1.5	719	6	BW245341	BW245341	C 928	18	1.5	782	9	CR260806	Reverse s	CR260806
C 856	18	1.5	719	6	CD448120	EL01N0204	C 929	18	1.5	785	7	CO533064	3530_1_21	CO533064
C 857	18	1.5	719	7	CF764149	CF764149	C 930	18	1.5	785	8	BZ644991	OGCJ36TC	BZ644991
C 858	18	1.5	720	6	CB171214	KBRE602601	C 931	18	1.5	785	9	AG532285	Mus muscu	AG532285
C 859	18	1.5	720	6	CD444748	EL01N0443	C 932	18	1.5	788	9	CNS02WLS	Tetradon	AL217202
C 860	18	1.5	720	6	CD446324	EL01T0403	C 933	18	1.5	789	6	CD433407	EL01N0308	CN824329
C 861	18	1.5	720	7	CG651821	Eg_CWGRS	C 934	18	1.5	789	7	CN824329	Oa_sp1bn	CN824329
C 862	18	1.5	722	1	AJ817704	AJ817704	C 935	18	1.5	789	7	CO522860	3530_1_15	CO522860
C 863	18	1.5	723	6	CD445994	EL01T0206	C 936	18	1.5	790	6	CD437334	EL01N0372	CD437334
C 864	18	1.5	723	7	CO517962	3530_1_11	C 937	18	1.5	790	7	CF728863	UI-M-HE0-	CF728863
C 865	18	1.5	723	8	CC175718	ZMMBC029	C 938	18	1.5	790	7	CK288224	EST750946	CK288224
C 866	18	1.5	724	7	CK026693	AGENCOURT	C 939	18	1.5	790	9	CC585958	CH240_383	CC585958
C 867	18	1.5	725	7	CF762580	CF762580	C 940	18	1.5	791	6	CD438326	EL01N0511	CD438326
C 868	18	1.5	727	6	CD435859	EL01N0367	C 941	18	1.5	791	8	BZ394682	EINBR16TR	BZ394682
C 869	18	1.5	727	6	CD444706	EL01N0443	C 942	18	1.5	791	9	AG500743	Mus muscu	AG500743
C 870	18	1.5	727	7	CK835178	CK835178	C 943	18	1.5	792	4	BG564169	602590145	BG564169
C 871	18	1.5	727	7	CK974791	CK974791	C 944	18	1.5	797	6	CD437839	EL01N0505	CD437839
C 872	18	1.5	728	9	AG290116	Mus muscu	C 945	18	1.5	798	8	BZ060741	IKG39B05.	BZ060741
C 873	18	1.5	729	5	BW017929	BW017929	C 946	18	1.5	799	6	CD434701	EL01N0327	CD434701
C 874	18	1.5	729	7	CK848958	971790 BA	C 947	18	1.5	799	8	BZ085658	11Q01402.	BZ085658
C 875	18	1.5	729	7	CK946793	CK946793	C 948	18	1.5	801	8	BZ116286	CH230-522	BZ116286
C 876	18	1.5	729	7	CO532608	CO532608	C 949	18	1.5	802	6	CD438448	EL01N0513	CD438448
C 877	18	1.5	731	6	CD442508	CD442508	C 950	18	1.5	803	7	CN824060	Oa_sp1bn	CN824060
C 878	18	1.5	731	7	CN822361	CN822361	C 951	18	1.5	804	6	CD438583	EL01N0515	CD438583
C 879	18	1.5	732	7	CK960589	CK960589	C 952	18	1.5	804	7	CO090150	GR_Ba09P	CO090150
C 880	18	1.5	733	9	AG285405	AG285405	C 953	18	1.5	805	7	CO875608	BovGen_03	CO875608
C 881	18	1.5	734	7	CO523068	CO523068	C 954	18	1.5	808	6	CD438227	EL01N0510	CD438227
C 882	18	1.5	735	6	CD436606	CD436606	C 955	18	1.5	809	6	CD433802	EL01N0315	CD433802
C 883	18	1.5	738	7	CN821853	CN821853	C 956	18	1.5	809	6	CD434773	EL01N0328	CD434773
C 884	18	1.5	739	7	CO525144	CO525144	C 957	18	1.5	809	6	CD437050	EL01N0367	CD437050
C 885	18	1.5	739	7	CO525144	CO525144	C 958	18	1.5	810	6	CD439624	EL01N0527	CD439624
C 886	18	1.5	740	6	CD445918	CD445918	C 959	18	1.5	810	6	CD438146	EL01N0509	CD438146
C 887	18	1.5	741	9	AG497257	AG497257	C 960	18	1.5	811	7	CO202417	Oa_sp1bn	CO202417
C 888	18	1.5	742	8	AZ743679	RPCI-24-1	C 961	18	1.5	813	6	CB166995	BTU602600	CB166995
C 889	18	1.5	743	9	AG551405	AG551405	C 962	18	1.5	814	6	CD437107	EL01N0368	CD437107
C 890	18	1.5	745	6	CB167125	GMM602702	C 963	18	1.5	814	8	BZ138326	CH230-386	BZ138326
C 891	18	1.5	746	6	CB453923	CB453923	C 964	18	1.5	818	8	AQ749952	HS_5576_A	AQ749952
C 892	18	1.5	747	4	BG386940	60245130	C 965	18	1.5	819	7	CO226007	WS01028_B	CO226007
C 893	18	1.5	747	7	CF408321	CF408321	C 966	18	1.5	820	6	CD437826	EL01N0505	CD437826
C 894	18	1.5	750	9	AG536368	AG536368	C 967	18	1.5	822	8	BZ567435	P8C82-164	BZ567435
C 895	18	1.5	751	6	CD783307	CD783307	C 968	18	1.5	823	7	CR455816	P8C82-164	CR455816
C 896	18	1.5	752	7	CN823219	CN823219	C 969	18	1.5	823	8	BZ726696	Q3FAG06TC	BZ726696
C 897	18	1.5	753	9	CR071955	CR071955	C 970	18	1.5	824	7	CO889966	BovGen_18	CO889966
C 898	18	1.5	756	6	CB317580	CB317580	C 971	18	1.5	825	6	CD434362	EL01N0323	CD434362
C 899	18	1.5	758	6	CD440350	CD440350	C 972	18	1.5	827	9	BX177180	Danio rer	BX177180
C 900	18	1.5	760	6	CB513936	CB513936	C 973	18	1.5	828	6	CD439608	EL01N0526	CD439608

Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1CM1862 row: a column: 16
High quality sequence stop: 595.
Location/Qualifiers
1. .787
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5094039"
/tissue_type="Burkitt lymphoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_8"
/notes="Organ: lymph; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
Query Match 1.9%; Score 22; DB 4; Length 787;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1157 AGCAATTTTACTTTTCTCTG 1178
|||||
Db 550 AGCAATTTTACTTTTCTCTG 529

RESULT 5
AK084319/c
LOCUS
DEFINITION Mus musculus 12 days embryo eyeball cDNA, RIKEN full-length enriched library, clone:D230023C08 product:unclassified, full insert sequence.
ACCESSION AK084319
VERSION AK084319.1 GI:26101882
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
10349636
REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
11042159
REFERENCE 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 888)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>.

FEATURES
source
1. .888
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D230023C08"
/db_xref="taxon:10090"
/clone="D230023C08"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
misc_feature 1. .888
/note="unclassifiable"

ORIGIN
Query Match 1.9%; Score 22; DB 3; Length 888;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 TTTAAACAATGTGATAAATAC 131
|||||
Db 467 TTTAAACAATGTGATAAATAC 446

RESULT 6
BF612800/c
LOCUS
DEFINITION BF612800 Wellcome CRC pcDNA1 egg xenopus laevis cDNA clone

```

IMAGE:3430450 3', mRNA sequence.
ACCESSION   BF612800
VERSION     BF612800.1  GI:11783795
KEYWORDS    EST.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM    Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
            Xenopodinae; Xenopus; Xenopus.
REFERENCE   1 (bases 1 to 174)
AUTHORS    Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L.,
            Pape,D., Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y.,
            Person,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
            WashU Xenopus EST project, 1999
            Unpublished (1999)
            Other ESTs: dd76906.y1
TITLE      WashU Xenopus EST project, 1999
JOURNAL     Unpublished (1999)
COMMENT     Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
            Gurdon, (Wellcome/CRC Institute). DNA Sequencing by: Washington
            University Genome Sequencing Center
            Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40UP from Gibco.
            Location/Qualifiers
FEATURES             source
    source            1..174
                     /organism="Xenopus laevis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:3430450"
                     /tissue_type="egg"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Wellcome CRC pcDNA1 egg"
                     /note="vector: pcDNA1; Site 1: NotI; Site 2: EcoRI; cDNAs
                     /note="were oligo-dT primed and directionally cloned. Library was
                     constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B.
                     Gurdon (Wellcome/CRC Institute)."

ORIGIN
Query Match      1.8%; Score 21; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 AAAAAAGCCAAATAAAAAA 34
    ||||||||||||||||||
Db 23 AAAAAAGCCAAATAAAAAA 3

RESULT 7
CL385624/c
LOCUS       CL385624 329 bp DNA linear GSS 19-AUG-2004
DEFINITION  RPCI44_332M17.f RPCI-44 Sus scrofa genomic clone RPCI44_332M17,
            genomic survey sequence.
ACCESSION   CL385624
VERSION     CL385624.1  GI:51437589
KEYWORDS    GSS.
SOURCE      Sus scrofa (pig)
ORGANISM    Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 329)
AUTHORS    Rogatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Marron,B.M.,
            Beaver,J.E. and Schook,L.B.
            Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
            Through Comparative Genomics
            Unpublished (2004)
            Other_GSSs: RPCI44_332M17.r

```

```

Contact: Lawrence B. Schook
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 265 5326
Fax: 217 244 5617
Email: schook@uiuc.edu
Clones are derived from the porcine BAC library RPCI-44
(http://www.bacpac.choi.org/porcine242.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@chori.org).
Clones may be purchased from BACPAC Resources
(http://BACPACResources.choi.org). This work was undertaken as part
of the International Swine Genome Sequencing Consortium by
University of Illinois at Urbana Champaign, USA with funds provided
by Grant No. AG2002-34480-11828 from USDA-CSREES and
AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
Initiative)
Plate: 332 row: M column: 17
Seq primer: T7
Class: BAC ends.
            Location/Qualifiers
FEATURES             source
    source            1..329
                     /organism="Sus scrofa"
                     /mol_type="genomic DNA"
                     /strain="four pigs (breed: 37.5% Yorks Landrace and 25%
                     Meishan)"
                     /db_xref="taxon:9823"
                     /clone="RPCI44_332M17"
                     /sex="male"
                     /cell_type="blood"
                     /clone_lib="RPCI-44"
                     /note="Vector: pTARBAC2; Site 1: EcoRI; Site 2: EcoRI;
                     porcine male BAC library produced by Pieter de Jong"

ORIGIN
Query Match      1.8%; Score 21; DB 9; Length 329;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1077 AGTCTGAGGATGAAGGAGG 1097
    ||||||||||||||||||
Db 320 AGTCTGAGGATGAAGGAGG 300

RESULT 8
BE478771/c
LOCUS       BE478771 390 bp mRNA linear EST 27-MAR-2003
DEFINITION  163162 BARC 5BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BE478771
VERSION     BE478771.1  GI:9598304
KEYWORDS    EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1 (bases 1 to 390)
AUTHORS    Sonstegard,T., Capuco,A.V., White,J., Van Tassell,C.P.,
            Connor,E.E., Cho,J., Sultana,R., Shade,L., Wray,J.E., Wells,K.D.
            and Quackenbush,J.
            Analysis of bovine mammary gland EST and functional annotation of
            the Bos taurus gene index
            Mamm. Genome 13 (7), 373-379 (2002)
            22135956
JOURNAL     12140684
MEDLINE
PUBMED
COMMENT     Contact: Sonstegard TS
            USDA, ARS, Beltsville Agricultural Research Center
            Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
            Tel: 301 504 8416
            Fax: 301 504 8414
            Email: tadsel@psi.barc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18

```

and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTACGACG
 Plate: 8 row: 1 column: 10
 Seq primer: ATTTAGGTGACACTATAG.
 Location/Qualifiers

FEATURES

source
 1. .390
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="BARC 5BOV"
 /note="vector; pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled mRNA isolated from mammary
 tissues at eight physiological, developmental, and disease
 states."

ORIGIN

Query Match 1.8%; Score 21; DB 2; Length 390;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1077 AGTCTGAGGATGGAAGGAGG 1097
 |||||
 Db 338 AGTCTGAGGATGGAAGGAGG 318

RESULT 9

AV668123 416 bp mRNA linear EST 28-NOV-2001
 LOCUS AV668123 Bos taurus ovary fetus Bos taurus cDNA clone EIOV019H11
 DEFINITION 3', mRNA sequence.
 ACCESSION AV668123
 VERSION AV668123.1 GI:9932868
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM

REFERENCE

AUTHORS Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
 and Sugimoto,Y.
 TITLE Establishment of a high throughput EST sequencing system using
 poly(A) tail-removed cDNA libraries and determination of 36,000
 bovine ESTs
 JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
 MEDLINE 21570554
 PUBMED 11713328
 COMMENT Contact: Yoshikazu Sugimoto
 Animal Genetics Division
 Shirakawa Institute of Animal Genetics
 Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
 Tel: 81-248-25-5641
 Fax: 81-248-25-5725
 Email: kazuugi@cocoa.ocn.ne.jp
 Single pass sequencing.
 This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source
 1. .416
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /clone="EIOV019H11"
 /tissue_type="ovary"
 /dev_stage="fetus"
 /lab_host="DH10B"
 /clone_lib="Bos taurus ovary fetus"
 /note="vector; pZLI; Site 1: SalI; Site 2: NotI; Poly A
 was deleted from a NotI site"

ORIGIN

Query Match 1.8%; Score 21; DB 1; Length 416;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1077 AGTCTGAGGATGGAAGGAGG 1097
 |||||
 Db 223 AGTCTGAGGATGGAAGGAGG 243

RESULT 10

AV596235 430 bp mRNA linear EST 27-NOV-2001
 LOCUS AV596235 Bos taurus cartilage fetus Bos taurus cDNA clone
 DEFINITION E1CA024E01 3', mRNA sequence.
 ACCESSION AV596235
 VERSION AV596235.1 GI:9713279
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM

REFERENCE

AUTHORS Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
 and Sugimoto,Y.
 TITLE poly(A) tail-removed cDNA libraries and determination of 36,000
 bovine ESTs
 JOURNAL Nucleic Acids Res. 29 (22), E108 (2001)
 MEDLINE 21570554
 PUBMED 11713328
 COMMENT Contact: Yoshikazu Sugimoto
 Animal Genetics Division
 Shirakawa Institute of Animal Genetics
 Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
 Tel: 81-248-25-5641
 Fax: 81-248-25-5725
 Email: kazuugi@cocoa.ocn.ne.jp
 Single pass sequencing.
 This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source
 1. .430
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /clone="E1CA024E01"
 /tissue_type="cartilage"
 /dev_stage="fetus"
 /lab_host="DH10B"
 /clone_lib="Bos taurus cartilage fetus"
 /note="vector; pZLI; Site 1: SalI; Site 2: NotI; Poly A
 was deleted from a NotI site"

ORIGIN

Query Match 1.8%; Score 21; DB 1; Length 430;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1077 AGTCTGAGGATGGAAGGAGG 1097
 |||||
 Db 219 AGTCTGAGGATGGAAGGAGG 239

RESULT 11

AQ594469 441 bp DNA linear GSS 08-JUN-1999
 LOCUS AQ594469
 DEFINITION HS 2105 B1 G06 T7C CIT Approved Human Genomic Sperm Library D Homo
 sapiens genomic clone Plate=2105 Col=11 Row=N, genomic survey
 sequence.
 ACCESSION AQ594469
 VERSION AQ594469.1 GI:5026055
 KEYWORDS GSS.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 441)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and
 Kaller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and
 Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.htsc.washington.edu
 Plate: 2105 row: N column: 11
 Seq primer: T7
 Class: BAC ends
 High quality sequence stop: 441.
 Location/Qualifiers
 1..441
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clones="plate:2105 Col=11 Row=N"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelBAC11; BAC Clones in
 E-Coli DH10B"

Query Match 1.8%; Score 21; DB 8; Length 441;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 488 AGGCATGACCCACCACTGGC 508
 |||||
 Db 404 AGGCATGACCCACCACTGGC 424

RESULT 12
 AZ475198/c
 LOCUS AZ475198 529 bp DNA linear GSS 04-OCT-2000
 DEFINITION lM0293J19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 clone UUGC1M0293J19 F, genomic survey sequence.
 ACCESSION AZ475198
 VERSION AZ475198.1 GI:10633323
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 529)
 REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
 Reilly,M., Rose,R., Rose,R., Stokes,R., Tingey,A., von
 Niederhausern,A. and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah
 University of Utah
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606

Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0293 row: J column: 19
 Seq primer: CGTTGTAAAACGACGGCCAGT
 Class: plasmid ends
 High quality sequence stop: 529.
 Location/Qualifiers
 1..529
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0293J19"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (GI|4732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

ORIGIN

Query Match 1.8%; Score 21; DB 8; Length 529;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 899 AATGAATCACACGATATT 919
 |||||
 Db 361 AATGAATCACACGATATT 341

RESULT 13
 BB659888/c
 LOCUS BB659888 672 bp mRNA linear EST 26-OCT-2001
 DEFINITION BB659888 RIKEN full-length enriched, 13 days embryo lung Mus
 musculus cDNA clone D430021A12 5', mRNA sequence.
 ACCESSION BB659888
 VERSION BB659888.1 GI:16493709
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 672)
 REFERENCE Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
 Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
 Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
 TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 JOURNAL Unpublished (2001)
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.
 and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., and
 Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and
 Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for
 further details.

e mouse tissues.

Location/Qualifiers

1. .672

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="D430021A12"

/tissue_type="lung"

/dev_stage="13 days embryo"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 13 days embryo

lung"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGCGCGCCCACTCGAGTCTTTTCTTTTCTTTT 3'], cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. Second strand cDNA was prepared with the
 primer adaptor of sequence [5'
 GAGAGAGAGTTCGAGTTAATTAATTCCTCCCCCCCC 3']. cDNA
 was cleaved with BamHI and XhoI. Vector: a modified
 pBluescript KS(+) after bulk excision from Lambda FLC I."

ORIGIN

Query Match 1.8% Score 21; DB 2; Length 672;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AAAAAAGCCAAATRAAAAA 34

|||||

Db 472 AAAAAAGCCAAATRAAAAA 452

RESULT 14

AJ452665/c

LOCUS

AJ452665 riken1 Gallus gallus cDNA clone 31mi4r1, mRNA sequence. EST 22-APR-2002

ACCESSION

VERSION

AJ452665.1 GI:20262761

KEYWORDS

EST.

SOURCE

ORGANISM

Gallus gallus (chicken)
 Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Buerstedde, J.M.

Unpublished (2002)

Contact: Buerstedde JM

Cellular Immunology

Heinrich-Pette-Institute

Martinstr. 52, 20251 Hamburg, Germany

Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.

FEATURES

source

1..717

/organism="Gallus gallus"

/mol_type="mRNA"

/db_xref="taxon:9031"

/clone="31mi4r1"

/cell_type="bursal lymphocyte"

/dev_stage="2-3 weeks old"

/clone_lib="riken1"

/notes="CB inbred strain"

ORIGIN

Query Match 1.8% Score 21; DB 1; Length 717;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 620 GCGCCACACGACGCG 640

|||||

Db 248 GCGCCACACGACGCG 228

RESULT 15

BY761301/c

LOCUS

DEFINITION

Mus musculus (house mouse)

SOURCE

ORGANISM

Mus musculus

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Nikaide, I., Oshino, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C.,
 Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Bruscia, V.,
 Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A.,
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A.,
 Kurachin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R.,
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,
 Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M.,
 Verdel, C., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Akakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,

Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Kogere, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,

Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,

Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,

Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,

Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by David A. Hume (Depts. of Biochemistry

and Microbiology/Parasitology Institute for Molecular Bioscience

University of Queensland Brisbane, Q 4072 Australia) whose

assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

FEATURES

source

Location/Qualifiers

1. .727

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="G530008P05"

/tissue_type="bone marrow"

/cell_type="macrophage"

/clone_lib="RIKEN full-length enriched, bone marrow

macrophage"

ORIGIN

Query Match 1.8%; Score 21; DB 6; Length 727;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 AAAAAAAAAAATAAAAAA 34

|||||

Db 363 AAAAAAAAAAATAAAAAA 343

Search completed: June 4, 2005, 18:31:18

Job time : 4506 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 04:22:34 ; Search time 166.757 Seconds
(without alignments)
780.981 Million cell updates/sec

Title: US-09-674-277-21
Perfect score: 22
Sequence: 1 ccactgacgataagcggaac 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseqn_16Dec04:*
1: Geneseqn19808:*
2: Geneseqn19908:*
3: Geneseqn20008:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	22	3	Aaz36121 Primer de
2	22	100.0	1181	3	Aaz36102 Nucleic a
C 3	17.8	80.9	2725	13	AdS58527 Bacterial
4	17.2	78.2	3715	4	ABL22996 Drosophil
5	17.2	78.2	3772	4	ABL22998 Drosophil
6	17.2	78.2	3772	4	AAS57170 DNA encod
7	17.2	78.2	3772	10	AdC35896 Drosophil
8	17.2	78.2	6051	12	ADP98744 C. albica
9	17.2	78.2	8496	6	ABQ76621 C. albica
C 10	16.8	76.4	1840	2	AAQ65611 Feline zo
C 11	16.8	76.4	1840	2	AAV64793 Feline 2P
C 12	16.8	76.4	1840	2	Aaz22716 Feline zo
C 13	16.8	76.4	1840	3	Aaz33250 Feline zo
C 14	16.8	76.4	1840	3	Aaz35653 Feline zo
C 15	16.8	76.4	1840	3	Aaz46261 Feline oo
C 16	16.8	76.4	1840	3	Aaz37805 Feline zo
C 17	16.8	76.4	1982	3	AAC42584 Arabidops
C 18	16.8	76.4	3327	4	ABL05013 Drosophil
19	16.8	76.4	4039	2	AAV99968 Nucleotid
20	16.8	76.4	4617	2	AAV99972 Nucleotid

C 21	16.8	76.4	5749	4	ABL18270
C 22	16.8	76.4	5905	4	ABL05012
C 23	16.8	76.4	8114	4	ABL03787
C 24	16.8	76.4	18737	4	ABL03786
C 25	16.4	74.5	545	6	AAL53422
C 26	16.2	73.6	683	4	ABL03277
C 27	16.2	73.6	730	10	ADK59601
C 28	16.2	73.6	1005	3	AAC45825
C 29	16.2	73.6	1005	6	ABZ13014
C 30	16.2	73.6	1005	12	ADN72926
C 31	16.2	73.6	1053	3	AAC45272
C 32	16.2	73.6	1101	10	ABX07448
C 33	16.2	73.6	1104	8	ACA50068
C 34	16.2	73.6	1128	13	ADR93541
C 35	16.2	73.6	1899	9	ADAL5526
C 36	16.2	73.6	1699	10	ADB31770
C 37	16.2	73.6	1699	10	ADD30689
C 38	16.2	73.6	1699	12	ADI43484
C 39	16.2	73.6	1699	12	ADO03298
C 40	16.2	73.6	1699	12	ADO01592
C 41	16.2	73.6	1863	8	ACA24360
C 42	16.2	73.6	2164	13	ADS57796
C 43	16.2	73.6	2664	2	AAQ78644
C 44	16.2	73.6	5536	4	ABL03276
C 45	16.2	73.6	10157	4	AAS46233

ALIGNMENTS

RESULT 1
AAZ36121

ID AAZ36121 standard; DNA; 22 BP.

XX AAZ36121;

DT 11-FEB-2000 (first entry)

DE Primer derived from a nucleic acid sequence specific to EHEC.

XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

XX Escherichia coli.

PN WO9555908-A2.

XX WO9555908-A2.

PD 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PA Frechon DTM, Laure FC, Thierry D;

PI WPI; 2000-013443/01.

DR New nucleic acid containing sequences specific to enterohemorrhagic

XX Escherichia coli, particularly serotype O157:H7, used for detecting these

PT bacteria in food.

XX Claim 5; Page 27; 48pp; French.

PS AAZ36103-27 represent fragments derived from nucleic acid sequences

CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are

CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to

CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and

CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).

CC The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies

XX
SQ Sequence 22 BP; 8 A; 7 C; 5 G; 2 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 3; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.54; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATTAAGCGGAAC 22

Db 1 CCACCTGAACGATTAAGCGGAAC 22

RESULT 2

AAZ36102

ID AAZ36102 standard; DNA; 1181 BP.

XX

AC AAZ36102;

XX

DT 11-FEB-2000 (first entry)

XX

DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX

KW Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

KW enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.

XX

OS Escherichia coli.

XX

PN W0955908-A2.

XX

PD 04-NOV-1999.

XX

PF 27-APR-1999; 99WO-FR001000.

XX

PR 28-APR-1998; 98PR-00005329.

XX

PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX

PI Pzechon DTM, Laure FC, Thierry D;

XX

DR WPI; 2000-013443/01.

XX

PT New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.

XX

PS Claim 1; Fig 2; 48pp; French.

XX

CC The present sequence is specific to enterohemorrhagic Escherichia coli
CC (EHEC). The sequence associated with the presence of virulence factors
CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
CC 68% homology with the virK gene which codes for virulence proteins of
CC Shigella flexneri. The present sequence is of plasmid origin. Fragments
CC of the present sequence are used, as probes and primers, for detection of
CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
CC animal samples, foods or the environment. The fragments are also useful
CC for epidemiological studies

SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 22; DB 3; Length 1181;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATTAAGCGGAAC 22

Db 718 CCACCTGAACGATTAAGCGGAAC 739

RESULT 3

ADS58527/c

ID ADS58527 standard; cDNA; 2725 BP.

XX

AC ADS58527;

XX

DT 02-DEC-2004 (first entry)

XX

DE Bacterial polynucleotide #10514.

XX

KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.

OS Bacteria.

XX

PN US2003233675-A1.

XX

PD 18-DEC-2003.

XX

PF 20-FEB-2003; 2003US-00369493.

XX

PR 21-FEB-2002; 2002US-0360039P.

XX

PA (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.

PA (GOLD/) GOLDMAN B S.

XX

PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX

WPI; 2004-061375/06.

XX

PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.

XX

PS Claim 1; SEQ ID NO 34201; 122pp; English.

XX

CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 2725 BP; 477 A; 909 C; 866 G; 473 T; 0 U; 0 Other;

```
Query Match      80.9%; Score 17.8; DB 13; Length 2725;
Best Local Similarity 90.5%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  2  CACCTGAACGATAAGCGGAAC 22
    |||||
Db   63  CACCGGACGATTAGCGGAAC 43

RESULT 4
ABL22996
ID  ABL22996 standard; DNA; 3715 BP.
XX
AC  ABL22996;
XX
DT  26-MAR-2002 (first entry)
XX
DE  Drosophila melanogaster genomic polynucleotide SEQ ID NO 20461.
XX
KW  Drosophila; developmental biology; cell signalling; insecticide;
KW  pharmaceutical; gene; ds.
XX
OS  Drosophila melanogaster.
XX
PN  WO200171042-A2.
XX
PD  27-SEP-2001.
XX
PF  23-MAR-2001; 2001WO-US009231.
XX
PR  23-MAR-2000; 2000US-0191637P.
PR  11-JUL-2000; 2000US-00614150.
XX
PA  (PEKE ) PE CORP NY.
XX
PI  Venter JC, Adams M, Li PWD, Myers EW;
XX
WI  WI; 2001-656860/75.
XX
PT  New isolated nucleic acid detection reagent for detecting 1000 or more
PT  genes from Drosophila and for elucidating cell signalling and cell-cell
PT  interactions.
XX
PS  Claim 1; SEQ ID NO 20461; 21pp + Sequence Listing; English.
XX
CC  The invention relates to an isolated nucleic acid detection reagent
CC  capable of detecting 1000 or more genes from Drosophila. The invention is
CC  useful in developmental biology and in elucidating cell signalling and
CC  cell-cell interactions in higher eukaryotes for the development of
CC  insecticides, therapeutics and pharmaceutical drugs. The invention
CC  discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC  sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC  ABB72072). The sequence data for this patent did not form part of the
CC  printed specification, but was obtained in electronic format directly
CC  from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 3715 BP; 927 A; 954 C; 943 G; 891 T; 0 U; 0 Other;

Query Match      78.2%; Score 17.2; DB 4; Length 3715;
Best Local Similarity 86.4%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  1  CCACCTGAACGATAAGCGGAAC 22
    |||||
Db   3521  CAATCTGAACGATAAGCGGCAC 3542

RESULT 5
ABL22998
ID  ABL22998 standard; DNA; 3772 BP.
XX
AC  ABL22998;
XX
```

```
DT  26-MAR-2002 (first entry)
XX
DE  Drosophila melanogaster genomic polynucleotide SEQ ID NO 20467.
XX
KW  Drosophila; developmental biology; cell signalling; insecticide;
KW  pharmaceutical; gene; ds.
XX
OS  Drosophila melanogaster.
XX
PN  WO200171042-A2.
XX
PD  27-SEP-2001.
XX
PF  23-MAR-2001; 2001WO-US009231.
XX
PR  23-MAR-2000; 2000US-0191637P.
PR  11-JUL-2000; 2000US-00614150.
XX
PA  (PEKE ) PE CORP NY.
XX
PI  Venter JC, Adams M, Li PWD, Myers EW;
XX
WI  WI; 2001-656860/75.
XX
PT  New isolated nucleic acid detection reagent for detecting 1000 or more
PT  genes from Drosophila and for elucidating cell signalling and cell-cell
PT  interactions.
XX
PS  Claim 1; SEQ ID NO 20467; 21pp + Sequence Listing; English.
XX
CC  The invention relates to an isolated nucleic acid detection reagent
CC  capable of detecting 1000 or more genes from Drosophila. The invention is
CC  useful in developmental biology and in elucidating cell signalling and
CC  cell-cell interactions in higher eukaryotes for the development of
CC  insecticides, therapeutics and pharmaceutical drugs. The invention
CC  discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC  sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC  ABB72072). The sequence data for this patent did not form part of the
CC  printed specification, but was obtained in electronic format directly
CC  from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 3772 BP; 933 A; 943 C; 931 G; 965 T; 0 U; 0 Other;

Query Match      78.2%; Score 17.2; DB 4; Length 3772;
Best Local Similarity 86.4%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  1  CCACCTGAACGATAAGCGGAAC 22
    |||||
Db   845  CAATCTGAACGATAAGCGGCAC 866

RESULT 6
AAS57170
ID  AAS57170 standard; DNA; 3772 BP.
XX
AC  AAS57170;
XX
DT  16-JAN-2002 (first entry)
XX
DE  DNA encoding Drosophila G-protein coupled receptor, GCPR #50.
XX
KW  Drosophila; G-protein coupled receptor; GCPR; insecticide; diagnostic;
KW  mutation detection; ds.
XX
OS  Drosophila melanogaster.
XX
PN  WO200170980-A2.
XX
PD  27-SEP-2001.
XX
PF  23-MAR-2001; 2001WO-US009341.
```

```
PR 23-MAR-2000; 2000US-0191638P.
XX 18-JUL-2000; 2000US-00618893.
XX (PEKE ) PE CORP NY.
XX Cravchik A;
XX WPI; 2001-616405/71.
XX P-PSDB; AAU38972.
XX
XX Sixty six Drosophila Melanogaster G-protein coupled receptors (GPCR),
PT useful in the treatment and diagnosis of GPCR-related conditions and for
PT identifying GPCR modulators for use as insecticides.
XX
XX Claim 4; Page 228-229; 392pp; English.
XX
XX The invention relates to sixty six novel isolated Drosophila melanogaster
CC G-protein coupled receptors (GPCR). The GPCR proteins and nucleic acids
CC are useful in the treatment and diagnosis of GPCR-related conditions. The
CC GPCR proteins and nucleic acids are also useful for identifying
CC modulators of GPCR proteins for use as insecticides. The nucleic acid can
CC also be used to detect mutations in GPCR genes and gene expression
CC products such as mRNA. AAS57072-AAS57203 represent D. melanogaster G-
CC coupled protein receptor genomic and coding sequences of the invention
XX
XX Sequence 3772 BP; 933 A; 943 C; 931 G; 965 T; 0 U; 0 Other;
SQ
Query Match 78.2%; Score 17.2; DB 4; Length 3772;
Best Local Similarity 86.4%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCACCTGAACGATAAGCGGAAC 22
Db 845 CAATCTGAACGATAAGCGGCAC 866
RESULT 7
ADC35896
ID ADC35896 standard; DNA; 3772 BP.
XX
XX ADC35896;
XX
XX 18-DEC-2003 (first entry)
XX
XX Drosophila G protein coupled receptor genomic DNA seq id 50.
XX
XX G-protein coupled receptor; GPCR; insecticide; drug screening;
XX insecticide screening; insecticidal activity; insecticidal tolerance;
XX fruit fly; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX US2003092124-A1.
XX
XX 15-MAY-2003.
XX
XX 15-OCT-2002; 2002US-00270333.
XX
XX 03-DEC-1999; 99US-0168677P.
XX 12-JAN-2000; 2000US-0175691P.
XX 23-MAR-2000; 2000US-0191638P.
XX 18-JUL-2000; 2000US-00618893.
XX
XX (APPL-) APPLERA CORP.
XX
XX Cravchik A;
XX
XX WPI; 2003-765480/72.
XX P-PSDB; ADC35898.
XX
XX New isolated G-protein coupled receptor useful for identifying modulators
PT as potential insecticides, to determine the biological activity of the
PT protein and for identifying compounds that modulate receptor activity.
PT
```

```
XX Claim 4; SEQ ID NO 148; 130pp; English.
XX
XX The invention describes an isolated protein (I) consisting or comprising
CC an amino acid sequence selected from fully defined 66 G-protein coupled
CC receptor amino acid sequences (S1), as given in the specification, an
CC allelic variant of (S1), an orthologue of (S1) or fragment of (S1). (I)
CC is useful for identifying an agent that binds to (I) which comprises
CC contacting the protein with an agent and assaying the contacted mixture
CC to determine whether a complex is formed with the agent bound to the
CC protein. (I) is useful for identifying modulators as potential
CC insecticides, to determine the biological activity of the protein (a
CC panel of multiple proteins for high-throughput screening), as targets for
CC identifying agents for use in human drugs and for identifying compounds
CC that modulate receptor activity. An antibody (II) that selectively binds
CC to (I) is useful for assessing normal and aberrant subcellular
CC localisation of cells and monitoring a treatment modality. A nucleic acid
CC (III) encoding (I) is useful for drug/insecticide screening to identify
CC compounds that modulate G-protein coupled receptor (GPCR) nucleic acid
CC expression, diagnostic assays for qualitative changes in GPCR nucleic
CC acid that lead to insecticidal activity/tolerance, to detect mutations in
CC GPCR genes and gene expression products such as mRNA, and as
CC hybridisation probes for determining the presence, level, form and
CC distribution of nucleic acid expression. A host cell comprising a vector
CC containing (III) is useful for conducting cell-based assays involving the
CC GPCR protein or its fragments, and identifying GPCR protein mutants. This
CC sequence encodes a fruit fly G-protein coupled receptor (GPCR).
XX
XX Sequence 3772 BP; 933 A; 943 C; 931 G; 965 T; 0 U; 0 Other;
SQ
Query Match 78.2%; Score 17.2; DB 10; Length 3772;
Best Local Similarity 86.4%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCACCTGAACGATAAGCGGAAC 22
Db 845 CAATCTGAACGATAAGCGGCAC 866
RESULT 8
ADP98744
ID ADP98744 standard; DNA; 6051 BP.
XX
XX ADP98744;
XX
XX 23-SEP-2004 (first entry)
XX
XX C. albicans specific gene, orfi9.1084, DNA sequence.
XX
XX Diploid fungal cell; allele; gene disruption cassette;
XX promoter replacement fragment; antifungal; fungicide; gene therapy;
XX infection; Candida albicans; gene; ds.
XX
XX Candida albicans.
XX
XX WO2004056965-A2.
XX
XX 08-JUL-2004.
XX
XX 19-DEC-2003; 2003WO-US040618.
XX
XX 19-DEC-2002; 2002US-0434832P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX (ELIT-) ELITRA CANADA LTD.
XX
XX Roemer T, Jiang B, Boone C, Bussey H;
XX WPI; 2004-500296/47.
XX P-PSDB; ADP99054.
XX
XX Constructing a strain of diploid fungal cells in which both alleles of a
PT gene are modified comprises modifying the alleles of a gene in the fungal
PT
```


PT cells by recombination using a gene disruption cassette and a promoter
XX replacement fragment.
XX
XX Claim 33; SEQ ID NO 6229; 163pp; English.
XX
XX The invention relates to a novel method for constructing a strain of
CC diploid fungal cells in which both alleles of a gene are modified. The
CC method comprises modifying the alleles of a gene in diploid fungal cells
CC by recombination using a gene disruption cassette and a promoter
CC replacement fragment. The invention further comprises: assembling a
CC collection of diploid fungal cells each of which comprises modified
CC alleles of a different gene; a strain of diploid fungal cells comprising
CC modified alleles of a gene, where the first allele of the gene is
CC inactivated by a gene disruption cassette comprising a nucleotide
CC sequence encoding an expressible selectable marker; and the expression of
CC the second allele of the gene is regulated by a heterologous promoter
CC that is operably linked to the coding region of the second allele of the
CC gene, and where the gene encodes the polypeptide mentioned above; a
CC collection of diploid fungal strains comprising the diploid strains cited
CC above, where substantially all the different genes that encode the above
CC amino acid sequences are modified and are present in different diploid
CC strains in the collection; a nucleic acid molecule microarray comprising
CC nucleic acid molecules, where each nucleic acid molecule comprises a
CC nucleotide sequence that is hybridizable to a target nucleotide sequence
CC comprising any of the 310 nucleotide sequences listed in the
CC specification (ADP98516-ADP98825); identifying a gene that is essential
CC to the survival or growth of a fungus, that contributes to the virulence
CC and/or pathogenicity of a fungus, or that contributes to the resistance
CC of a diploid fungus to an antifungal agent; identifying an antifungal
CC agent that inhibits the growth of a diploid fungus, or a therapeutic
CC agent for treatment of a mammalian disease; correlating changes in the
CC levels of proteins or gene transcripts with the inhibition of growth or
CC proliferation of a diploid fungal cell; a purified or isolated nucleic
CC acid molecule comprising a nucleotide sequence encoding a gene product
CC required for proliferation of *Candida albicans*, where the gene product
CC consists of any of the above-mentioned amino acid sequences; a vector
CC comprising a promoter operably linked to the nucleic acid molecule cited
CC above; a host cell containing the vector; a purified or isolated
CC polypeptide comprising any of the 61 amino acid sequences given in the
CC specification (ADP96718-ADP96778); a fusion protein comprising a fragment
CC of a first polypeptide fused to a second polypeptide, the fragment
CC consisting of at least 6 consecutive residues of any of ADP98826-ADP99135
CC; producing a polypeptide; identifying a compound which modulates the
CC activity of a gene product encoded by a nucleic acid comprising any of
CC ADP98516-ADP98825; eliciting an immune response in an animal; a strain of
CC *Candida albicans*, where a first allele of a gene comprising any of
CC ADP98516-ADP98825 is inactive and a second allele of the gene is under
CC the control of a heterologous promoter; identifying a compound or binding
CC partner that binds to the polypeptide comprising any of ADP98826-
CC ADP99135, or its fragment; identifying a compound having the ability to
CC inhibit growth or proliferation of *Candida albicans*; inhibiting growth or
CC proliferation of *Candida albicans* cells; manufacturing an antimycotic
CC compound; treating an infection of a subject by *Candida albicans*;
CC preventing or containing contamination of an object by *Candida albicans*,
CC or for preventing or inhibiting formation on a surface of a biofilm
CC comprising *Candida albicans*; a pharmaceutical composition comprising a
CC therapeutic amount of an agent which reduces the activity or level of a
CC gene product encoded by a nucleic acid comprising any of ADP98516-
CC ADP98825 in a pharmaceutical carrier; an antibody preparation which binds
CC the polypeptide; methods for evaluating a compound against a target gene
CC product encoded by any of ADP98516-ADP98825; identifying an antimycotic
CC compound; a computer or a computer readable medium that comprises at
CC least one of the nucleotide sequences mentioned in the specification or
CC at least one amino acid sequence selected from ADP98826-ADP99135; a
CC method assisted by a computer for identifying a putatively essential gene
CC of a fungus; and a protein array comprising proteins, where at least one
CC sequence comprises an amino acid sequence of a portion of an amino acid
CC sequence selected from ADP98516-ADP98825. The novel methods and
CC compositions have fungicide activity. The compositions may be used in
CC gene therapy. The composition and methods are useful for drug screening
CC purposes or for diagnosing, preventing or treating infections associated
CC with *Candida albicans*. These may also be used for constructing strains
CC useful for identification and validation of gene products as effective

CC targets for therapeutic intervention, for identifying and validating gene
CC products as effective targets for therapeutic intervention, and for
CC collecting identified essential genes. This polynucleotide sequence
CC represents a *Candida albicans* fungal specific gene of the invention.
CC NOTE: This sequence was downloaded from an electronic sequence listing
CC provided on the WIPO website.
XX
XX Sequence 6051 BP; 2049 A; 1041 C; 1129 G; 1832 T; 0 U; 0 Other;
SQ
Query Match 78.2%; Score 17.2; DB 12; Length 6051;
Best Local Similarity 86.4%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCACCTGAACGATACGCGGAAC 22
Db 2857 CCACCTCAACACACGCGGAAC 2878
RESULT 9
ABQ76621
ID ABQ76621 standard; cDNA; 8496 BP.
XX
XX ABQ76621;
XX
XX 21-NOV-2002 (first entry)
XX
XX C. *albicans* BAX-associated cDNA fragment SEQ ID 667.
XX
XX Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
XX vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
XX apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
XX neurodegeneration; cell death; ss.
XX
XX *Candida albicans*.
XX
XX WO200264766-A2.
XX
XX 22-AUG-2002.
XX
XX 21-DEC-2001; 2001WO-EP015398.
XX
XX 22-DEC-2000; 2000EP-00870318.
XX
XX 04-JAN-2001; 2001EP-00870002.
XX
XX 09-JAN-2001; 2001EP-00870003.
XX
XX (JANC) JANSSEN PHARM NV.
XX
XX Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
XX WPI; 2002-667002/71.
XX
XX P-PSDB; ABG93355.
XX
XX New isolated nucleic acid representing a synthetic BAX-gene, useful as
XX medicament for treating, preventing and/or alleviating yeast or fungal
XX infections or proliferative disorders, or for preventing apoptosis in
XX certain diseases.
XX
XX Claim 36; Fig 2; 344pp; English.
XX
XX This invention describes a novel nucleic acid representing a synthetic
XX Bax gene. The Bax gene of the invention is useful for identifying Bax-
XX resistant yeast or fungi, identifying, or obtaining and identifying
XX *Candida spp.* sequences that are differentially expressed in a pathway
XX eventually leading to programmed cell death or identifying inhibitors or
XX inhibitor sequences of Bax-induced cell death. The products of the
XX invention have cytostatic, fungicide; immunosuppressive, virucide and
XX vasotropic activity and can be used in vaccines or for gene therapy. The
XX isolated nucleic acids, polypeptides, pharmaceutical compositions,
XX antisense molecules and antibodies are useful as medicaments or in
XX preparing a medicament for treating, preventing and/or alleviating
XX diseases associated with yeast or fungi or proliferative disorders, such
XX as cancer, or for preventing apoptosis in certain diseases. The compounds
XX or polypeptides, or the genetically modified organism are useful for

CC preparing a medicament for modifying the endogenic flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polynucleotide associated with the Bax gene
 CC described in the disclosure of the invention

XX
 SQ Sequence 8496 BP; 2931 A; 1542 C; 1387 G; 2636 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 6; Length 8496;
 Best Local Similarity 86.4%; Pred. No. 2.4e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
 ||||| ||||| ||||| |||||
 Db 6726 CCACCTCAACACAAAGCGGAAC 6747

RESULT 10
 AAQ65611/c
 ID AAQ65611 standard; cDNA; 1840 BP.

XX AAQ65611;
 XX
 DT 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 01-FEB-1995 (first entry)

DE Feline zona pellucida ZPB coding sequence.

XX Cat; feline; zona pellucida; ZPB; immunocontraception; ds.

XX Felis catus.

PH Key Location/Qualifiers
 FT CDS 57..1769
 FT /*tag= a
 FT /product= "ZPB"

XX WO9411019-A1.

XX 26-MAY-1994.

XX 06-NOV-1993; 93WO-US010851.

XX 09-NOV-1992; 92US-00973341.

XX 29-JAN-1993; 93US-00012990.

XX (ZONA-) ZONAGEN INC.

XX Harris JD, Hsu KT, Podolski JS;

XX WPI; 1994-183156/22.

XX P-PSDB; AAR55201.

XX Use of zona pellucida proteins and antibodies - for inducing reproducible
 transient infertility or permanent sterility in female mammals.

XX Claim 24; Page 99-101; 154pp; English.

XX A cDNA library was prepared in lambda gt10 from mRNA isolated from
 ovaries of 3-4 month old cats. Plaques were screened using a mixture of
 CC probes encoding porcine ZPA, ZPB and ZPC proteins. Positive clones were
 CC analysed further by Southern hybridisation using the porcine probes and
 CC clones encoding feline ZPA, ZPB and ZPC proteins were identified. The
 CC deduced amino acid sequence from the feline ZPB clone was approximately
 CC 80% homologous to porcine ZPB protein. (Updated on 25-MAR-2003 to correct
 CC PN field.) (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 2; Length 1840;
 Best Local Similarity 90.0%; Pred. No. 3.2e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAGCGGA 21
 ||||| ||||| ||||| |||||
 Db 1061 CACCTGAACGCTAAGCGGA 1042

RESULT 11
 AAV64793/c
 ID AAV64793 standard; cDNA; 1840 BP.

XX AAV64793;

XX 29-JAN-1999 (first entry)

XX Feline ZPB cDNA.

XX ZPB; zona pellucida; infertility; sterility; immunocontraceptive;
 KW vaccine; feline; ds.

XX Felis sp.

PH Key Location/Qualifiers
 FT CDS 57..1769
 FT /*tag= a
 FT /product= "ZPB"

XX US5837497-A.

XX 17-NOV-1998.

XX 07-JUN-1995; 95US-00484993.

XX 09-NOV-1992; 92US-00973341.

XX 28-JAN-1993; 93US-00012990.

XX 09-NOV-1993; 93US-00149223.

XX (ZONA-) ZONAGEN INC.

XX Harris JD, Hsu KT, Podolski JS;

XX WPI; 1999-023447/02.

XX P-PSDB; AAW81811.

XX Isolated zona pellucida DNA from different mammals - used to develop
 products which can be used for vaccination to induce transient
 PT infertility or permanent sterility in female mammals.

XX Example 4; Col 89-92; 84pp; English.

XX This sequence encodes a feline ZPB protein isolated from zona pellucida.
 CC This protein can be used in a method for specifically inducing transient
 CC infertility or permanent sterility in a host animal by selective
 CC vaccination with specific zona pellucida proteins or
 CC immunocontraceptively active fragments

XX Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 2; Length 1840;

Best Local Similarity 90.0%; Pred. No. 3.2e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAGCGGA 21
 ||||| ||||| ||||| |||||
 Db 1061 CACCTGAACGCTAAGCGGA 1042

RESULT 12
 AAZ22716/c
 ID AAZ22716 standard; cDNA; 1840 BP.

XX AAZ22716;

XX

DT 17-OCT-2003 (revised)
 DT 15-FEB-2000 (first entry)
 XX Feline zona pellucida ZPB gene.
 DE
 DE
 XX
 KW Contraceptive; porcine; pig; zona pellucida; mammal; lapine; canine; dog;
 KW rabbit; feline; cat; bovine; cow; human; cynomolgus monkey; ZPA; ZPB;
 KW ZPC; infertility; veterinary; ds.
 XX
 XX Felis catus.
 OS
 XX
 XX US5981228-A.
 PN
 XX
 XX
 PD 09-NOV-1999.
 XX
 XX 07-JUN-1995; 95US-00484596.
 PF
 XX 09-NOV-1992; 92US-00973341.
 PR 29-JAN-1993; 93US-00012990.
 PR 09-NOV-1993; 93US-00149223.
 XX
 PA (ZONA-) ZONAGEN INC.
 XX
 XX Podolski JS, Hsu KT, Harris JD;
 PI
 XX
 XX WPI; 1999-633318/54.
 DR P-PSDB; AAY42474.
 DR
 XX Mammalian zona pellucida proteins used to induce transient or permanent
 PT infertility.
 PT
 XX
 XX Example 4; Col 87-92; 84pp; English.
 PS
 CC This sequence represents the coding region for the feline zona pellucida
 CC ZPB protein. The invention relates to the isolation of novel nucleotide
 CC sequences encoding zona pellucida proteins from mammalian, especially
 CC porcine, lapine, canine, feline, bovine, human or cynomolgus monkey
 CC sources. The zona pellucida proteins (ZPA, ZPB, and ZPC) are used in the
 CC induction of transient or permanent infertility. At present the method is
 CC used in veterinary applications to induce transient or permanent
 CC infertility in porcine, lapine, canine, feline, bovine, and cynomolgus
 CC monkeys. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T; 0 U; 0 Other;
 SQ
 Query Match 76.4%; Score 16.8; DB 2; Length 1840;
 Best Local Similarity 90.0%; Pred. No. 3.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 CACCTGAACGATAAGCGGAA 21
 DB 1061 CACCTGAACGCTAAGCGGGA 1042
 RESULT 13
 AAZ33250/c
 ID AAZ33250 standard; cDNA; 1840 BP.
 XX
 AC AAZ33250;
 XX
 DT 14-FEB-2000 (first entry)
 XX
 DE Feline zona pellucida ZPB nucleotide sequence.
 XX
 KW Zona pellucida; ZPA; ZPB; ZPC; infertility; sterility; vaccine;
 KW immunocontraceptive; contraception; immunisation; ds.
 XX
 XX Felis sp.
 OS
 XX US5976545-A.
 PN
 XX
 PD 02-NOV-1999.
 XX

PF 07-JUN-1995; 95US-00484158.
 XX
 PR 09-NOV-1992; 92US-00973341.
 PR 29-JAN-1993; 93US-00012990.
 PR 09-NOV-1993; 93US-00149223.
 XX
 PA (ZONA-) ZONAGEN INC.
 XX
 XX Podolski JS, Hsu KT, Harris JD;
 PI
 XX
 XX WPI; 2000-037080/03.
 DR P-PSDB; AAY52978.
 DR
 XX New recombinant mammalian zona pellucida C proteins, for induction of
 PT female sterility.
 PT
 XX
 XX Example 4; Col 87-92; 87pp; English.
 PS
 CC The present invention describes recombinant zona pellucida (ZP) proteins,
 CC specifically ZPC proteins. Also described are fusion proteins of ZPC with
 CC any of keyhole limpet haemocyanin, muramyl dipeptide, histidine tag, beta
 CC -galactosidase or palmitic acid, capable of stimulating mammalian
 CC antibodies that recognise mammalian ZPC. Recombinant ZPC proteins are
 CC used to immunise animals to induce transient infertility or permanent
 CC female sterility in humans or other mammals. AAZ33243 to AAZ33254,
 CC AAZ33270 to AAZ33274 and AAZ33285 encode mammalian ZP proteins given in
 CC AAY52971 to AAY52988 from the present invention. AAZ33255 to AAZ33269 and
 CC AAZ33275 to AAZ33284 represent oligonucleotides used in the
 CC exemplification of the present invention
 XX
 XX Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T; 0 U; 0 Other;
 SQ
 Query Match 76.4%; Score 16.8; DB 3; Length 1840;
 Best Local Similarity 90.0%; Pred. No. 3.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 CACCTGAACGATAAGCGGAA 21
 DB 1061 CACCTGAACGCTAAGCGGGA 1042
 RESULT 14
 AAZ95653/c
 ID AAZ95653 standard; cDNA; 1840 BP.
 XX
 AC AAZ95653;
 XX
 DT 15-SEP-2003 (revised)
 DT 12-JUN-2000 (first entry)
 XX
 DE Feline zona pellucida ZPB encoding cDNA SEQ ID NO:15.
 XX
 KW Zona pellucida; ZPA; ZPB; ZPC; infertility; permanent sterility; vaccine;
 KW immunocontraceptive; contraception; immunostimulatory; ds.
 XX
 XX Felis catus.
 OS
 XX
 XX Key Location/Qualifiers
 FT CDS 57..1769
 FT /tag= a
 FT /product= "ZPB"
 FT /note= "zona pellucida protein"
 XX
 XX US6027727-A.
 PN
 XX
 XX 22-FEB-2000.
 PD
 XX
 XX 09-NOV-1993; 93US-00149223.
 PF
 XX
 XX 09-NOV-1992; 92US-00973341.
 PR 29-JAN-1993; 93US-00012990.
 PR
 XX
 XX (ZONA-) ZONAGEN INC.

XX Podolski JS, Hsu KT, Harris JD;
 XX WPI; 2000-269144/23.
 DR P-PSDB; AAY82209.
 XX
 PT Inducing reproducible transient or permanent infertility in a mammal
 PT comprises administration of homologous and/or heterologous mammalian
 PT species zona pellucida proteins.
 XX
 PS Example 4; Col 89-94; 85pp; English.
 XX
 CC The present invention describes a method for inducing reproducible
 CC transient infertility in a female mammal, including humans, by selective
 CC administration of homologous and/or heterologous mammalian species zona
 CC pellucida (ZP) proteins or immunocontraceptively active fragments. Also
 CC described are: (1) a method for inducing transient infertility in a
 CC mammal by administering directly an antibody directed to a ZPB or an
 CC immunologically active fragment selected from feline, bovine, cynomolgous
 CC monkey or human ZPB polypeptides; (2) an isolated, purified recombinant
 CC feline, bovine, cynomolgous monkey or human ZPB polypeptide or
 CC immunologically active fragment; and (3) a fusion protein comprising a
 CC ZPB polypeptide which is conjugated with a compound selected from keyhole
 CC limpet haemocyanin, muramyl dipeptide, histidine-tag, beta-gal, or
 CC palmitic acid where the fusion protein remains effective to stimulate
 CC production of antibodies that recognise a ZPB polypeptide. The method is
 CC useful for providing transient or permanent infertility or sterility in
 CC humans and animals. The present sequence encodes feline ZPB, which is
 CC used in the exemplification of the present invention. (Updated on 15-SEP-
 CC 2003 to standardise OS field)

XX SQ Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 3; Length 1840;
 Best Local Similarity 90.0%; Pred. No. 3.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAA 21
 |||||
 DB 1061 CACCTGAACGCTAAGCGGGA 1042

RESULT 15
 AAZ46261/c
 ID AAZ46261 standard; cDNA; 1840 BP.
 XX
 AC AAZ46261;
 XX
 DT 15-SEP-2003 (revised)
 DT 14-MAR-2000 (first entry)
 XX
 DE Feline oocyte zona pellucida protein ZPB cDNA.
 XX
 KW Zona pellucida; ZPA; ZPB; ZPC; purified; mammalian; glycoprotein; target;
 KW immunocontraception; vaccine; antibody; transient; infertility;
 KW controllable; predictable; permanent; sterility; side effect; ds.
 XX
 OS Felis catus.

Key Location/Qualifiers
 CDS 57..1769
 FT /*tag= a
 FT /product= "Feline oocyte zona pellucida protein ZPB"

US6001599-A.
 14-DEC-1999.
 02-JUN-1995; 95US-00458731.
 09-NOV-1992; 92US-00973341.
 29-JAN-1993; 93US-00012990.
 09-NOV-1993; 93US-00149223.

XX (ZONA-) ZONAGEN INC.
 XX Podolski JS, Hsu KT, Harris JD;
 XX WPI; 2000-061880/05.
 XX
 PT Isolated DNA encoding mammalian zona pellucida proteins useful for
 PT inducing transient sterility.
 XX
 PS Claim 1; Col 87-92; 84pp; English.
 XX
 CC This sequence represents cDNA encoding a mammalian zona pellucida
 CC protein, ZPB. The zona pellucida is a complex matrix surrounding the
 CC mammalian oocyte, formed of glycoproteins secreted by ovarian cells. The
 CC invention relates to zona pellucida proteins ZPA, ZPB and ZPC, which are
 CC useful as targets for immunocontraceptive vaccines. Polynucleotides
 CC encoding these proteins can be used for recombinant protein production,
 CC and as probes in hybridisation methods for the isolation of
 CC polynucleotides encoding zona pellucida proteins from other mammalian
 CC species. Administration of zona pellucida proteins to a host animal
 CC results in a specific immunocontraceptive effect. Administration of
 CC purified ZPA and/or ZPB induces transient infertility in females.
 CC Fertility can be maintained in an on or off condition in a controllable/
 CC predictable fashion. Administration of ZPC induces permanent sterility.
 CC Use of a purified zona pellucida protein rather than a complex mixture of
 CC zona pellucida proteins reduces the potential for unwanted side-effects
 CC which may be harmful. (Updated on 15-SEP-2003 to standardise OS field)

XX SQ Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 3; Length 1840;
 Best Local Similarity 90.0%; Pred. No. 3.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAA 21
 |||||
 DB 1061 CACCTGAACGCTAAGCGGGA 1042

RESULT 16
 AAZ37805/c
 ID AAZ37805 standard; DNA; 1840 BP.
 XX
 AC AAZ37805;
 XX
 DT 15-SEP-2003 (revised)
 DT 09-FEB-2000 (first entry)
 XX
 DE Feline zona pellucida protein B (ZPB) gene.
 XX
 KW Zona pellucida protein; ZPA; ZPB; ZPC; oocyte; ovarian cell; antibody;
 KW fusion protein; vaccine; contraceptive; fertilization; fertility;
 KW immunocontraceptive; ss.

OS Felis catus.
 Key Location/Qualifiers
 CDS 57..1769
 FT /*tag= a
 FT /product= "ZPB"
 FT /notes= "Zona pellucida protein B"

US5989550-A.
 23-NOV-1999.
 07-JUN-1995; 95US-00480150.
 09-NOV-1992; 92US-00973341.
 29-JAN-1993; 93US-00012990.
 09-NOV-1993; 93US-00149223.

PA (ZONA-) ZONAGEN INC.
 XX
 PI Podolski JS, Hsu KT, Harris JD;
 XX
 XX WPI; 2000-022762/02.
 DR P-PSDB; AAY52174.
 DR
 XX Isolated and purified mammalian zona pellucida proteins useful in methods
 PT of immunoreception.
 PT
 XX
 XX Example 4; Col 87-92; 84pp; English.
 XX
 CC Sequences AAZ37804-Z37806 are feline zona pellucida proteins ZPA, ZPB,
 CC and ZPC genes. The invention relates to isolated and purified zona
 CC pellucida proteins from mammals, specifically canine, feline and bovine
 CC ZPA or their biologically active fragments also porcine and cynomolgus
 CC monkey ZPA. The zona pellucida is a complex matrix surrounding the
 CC mammalian oocyte, formed of glycoproteins secreted by ovarian cells. The
 CC previously mentioned ZPA proteins can be used to produce a fusion protein
 CC to stimulate production of antibodies in a mammal that recognize a ZPA
 CC polypeptide of the mammal. The ZPA polypeptides are useful as vaccines
 CC for selectively effecting transient infertility in mammals. ZP has an
 CC important role in maintaining the oocyte and in oocyte-sperm interactions
 CC and so is a target for a contraceptive agent which interferes with the
 CC fertilization process. Providing a specific polypeptide reduces the need
 CC for a complex mixture of many ZP proteins which may create unwanted and
 CC harmful side effects. The duration of transient fertility is controllable
 CC and can be maintained in a controllable and/or predictable fashion.
 CC (Updated on 15-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 1840 BP; 433 A; 538 C; 441 G; 428 T; 0 U; 0 Other;
 76.4%; Score 16.8; DB 3; Length 1840;
 Query Match Best Local Similarity 90.0%; Pred. No. 3.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 CACCTGAACGATTAAGCGGAA 21
 |||||
 Db 1061 CACCTGAACGATTAAGCGGGA 1042
 RESULT 17
 AAC42584/c
 ID AAC42584 standard; DNA; 1982 BP.
 XX
 AC AAC42584;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36101.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway; metabolic pathway;
 KW promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-00301439.
 XX
 XX 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 99US-0130077P.
 99US-0130449P.
 99US-0130510P.
 99US-0130891P.
 99US-0131449P.
 99US-0132048P.
 99US-0132407P.
 99US-0132484P.
 99US-0132485P.
 99US-0132486P.
 99US-0132487P.
 99US-0132863P.
 99US-0134256P.
 99US-0134218P.
 99US-0134219P.
 99US-0134221P.
 99US-0134370P.
 99US-0134370P.
 99US-0134768P.
 99US-0134941P.
 99US-0135124P.
 99US-0135353P.
 99US-0135629P.
 99US-0136021P.
 99US-0136392P.
 99US-0136782P.
 99US-0137222P.
 99US-0137528P.
 99US-0137502P.
 99US-0137724P.
 99US-0138094P.
 99US-0138540P.
 99US-0138847P.
 99US-0139119P.
 99US-0139452P.
 99US-0139453P.
 99US-0139492P.
 99US-0139454P.
 99US-0139455P.
 99US-0139456P.
 99US-0139457P.
 99US-0139458P.
 99US-0139459P.
 99US-0139460P.
 99US-0139461P.
 99US-0139462P.
 99US-0139463P.
 99US-0139750P.
 99US-0139763P.
 99US-0139817P.
 99US-0139899P.
 99US-0140353P.
 99US-0140354P.
 99US-0140695P.
 99US-0140823P.
 99US-0140991P.
 99US-014287P.
 99US-0141842P.
 99US-0142154P.
 99US-0142055P.
 99US-0142390P.
 99US-0142803P.
 99US-0142920P.
 99US-0142977P.
 99US-0143542P.
 99US-0143624P.
 99US-0144005P.
 99US-0144085P.
 99US-0144086P.
 99US-0144325P.
 99US-0144331P.
 99US-0144332P.
 99US-0144333P.
 99US-0144334P.

PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 76.4%; Score 16.8; DB 3; Length 1982;

Best Local Similarity 90.0%; Pred. No. 3.2e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAAGCGGAAC 22

||||||| |||||

Db 877 ACCTGAACAATAATCGGAAC 858

RESULT 18

ABL05013

ID ABL05013 standard; cDNA; 3327 BP.

XX ABL05013;

XX AC ABL05013;

XX DT 26-MAR-2002 (first entry)

XX XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 9521.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical; gene; ss.

XX OS Drosophila melanogaster.

XX OS WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX DR P-PSDB; ABB60910.

XX XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 9521; 21pp + Sequence Listing; English.
PS The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
CC

CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-
 CC AB872072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 3327 BP; 923 A; 787 C; 794 G; 823 T; 0 U; 0 Other;
 Query Match 76.4%; Score 16.8; DB 4; Length 3327;
 Best Local Similarity 90.0%; Pred. No. 3.4e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAAGCGGAAC 22
 | | | | | | | | | | | | | | | | | | | | | |
 Db 581 AGCTGAACGATAAGCTGAAC 600

RESULT 19
 AAV99968
 ID AAV99968 standard; cDNA; 4039 BP.
 XX AC AAV99968;
 XX DT 07-JUN-1999 (first entry)
 XX DE Nucleotide sequence of the Drosophila smoothened gene.
 XX KW Drosophila; smoothened gene; smoothened peptides; SPs;
 KW cell proliferation; cell differentiation; spermatogenesis; osteogenesis;
 KW chondrogenesis; apoptosis; degenerative disorder; neoplastic disorder;
 KW hyperplastic disorder; varicella-zoster; tumour; liver repair; ds.
 XX OS Drosophila sp.
 XX FH Key Location/Qualifiers
 FT CDS 271..3381
 FT /*tag= a
 FT /product= "smoothened polypeptide"

XX
 PN W05901468-A2.
 XX PD 14-JAN-1999.
 XX PF 01-JUL-1998; 98WO-US013793.
 XX PR 01-JUL-1997; 97US-0051407P.
 XX PR 21-JUL-1997; 97US-00897798.
 XX PA (ONTO-) ONTOGENY INC.
 PA (INGH/) INGHAM P W.
 PA (VHEU/) VAN DEN HEUVEL M.
 XX WPI; 1999-105996/09.
 DR P-PSDB; AAW81063.
 XX
 XX New isolated smoothened genes - used to develop products for treating
 PT e.g. cancers, neurodegenerative disorders, nervous system injury,
 PT immunological diseases or infections.
 XX
 XX Claim 12; Page 87-89; 107pp; English.

XX This is the nucleotide sequence of the novel Drosophila smoothened gene
 CC used in the method of the invention to produce smoothened polypeptides
 CC (SPs). The SPs can modulate at least one of proliferation,
 CC differentiation or survival of a cell which expresses the SP. The
 CC products can be used to modulate spermatogenesis, osteogenesis,
 CC chondrogenesis or neuronal cell differentiation or to enhance survival of
 CC neuronal cells, e.g. to prevent apoptosis. They can be used for treating
 CC e.g. degenerative disorders marked by loss of particular cell-types,
 CC apoptosis, neoplastic and/or hyperplastic disorders. In particular they

CC can be used to treat neurological conditions derived from acute,
 CC subacute, or chronic injury to the nervous system, aging of the nervous
 CC system chronic neurodegenerative diseases of the nervous system, chronic
 CC immunological diseases of the nervous system or affecting the nervous
 CC system. They can also treat be used to CNS trauma infection, infection
 CC (such as viral infection with varicella-zoster), metabolic disease,
 CC nutritional deficiency, toxic agents (such as cisplatin treatment),
 CC chronic pain syndromes or in nerve prostheses for the repair of central
 CC and peripheral nerve damage, malignant gliomas, medulloblastomas,
 CC neuroectodermal tumours, ependymomas and for liver repair and
 CC regeneration of lung tissue in the treatment of emphysema, or for
 CC enhancing bond formation, tissue formation or fertility. The products can
 CC also be used for detection, diagnosis and drug screening

XX
 SQ Sequence 4039 BP; 1164 A; 910 C; 905 G; 1060 T; 0 U; 0 Other;
 Query Match 76.4%; Score 16.8; DB 2; Length 4039;
 Best Local Similarity 90.0%; Pred. No. 3.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAAGCGGAAC 22
 | | | | | | | | | | | | | | | | | | | | | |
 Db 635 AGCTGAACGATAAGCTGAAC 654

RESULT 20
 AAV99972
 ID AAV99972 standard; DNA; 4617 BP.
 XX AC AAV99972;
 XX DT 07-JUN-1999 (first entry)
 XX DE Nucleotide sequence of the smoothened gene.
 XX KW Smoothened gene; smoothened peptides; SPs; liver repair;
 KW cell proliferation; cell differentiation; spermatogenesis; osteogenesis;
 KW chondrogenesis; apoptosis; degenerative disorder; neoplastic disorder;
 KW hyperplastic disorder; varicella-zoster; tumour; ds.
 XX OS Synthetic.
 XX PN W09901468-A2.
 XX PD 14-JAN-1999.
 XX PF 01-JUL-1998; 98WO-US013793.
 XX PR 01-JUL-1997; 97US-0051407P.
 XX PR 21-JUL-1997; 97US-00897798.
 XX PA (ONTO-) ONTOGENY INC.
 PA (INGH/) INGHAM P W.
 PA (VHEU/) VAN DEN HEUVEL M.
 XX WPI; 1999-105996/09.
 XX
 XX New isolated smoothened genes - used to develop products for treating
 PT e.g. cancers, neurodegenerative disorders, nervous system injury,
 PT immunological diseases or infections.
 XX
 XX Claim 12; Page 105-107; 107pp; English.

XX This is the (genomic) nucleotide sequence of the novel smoothened gene
 CC used in the method of the invention to produce smoothened polypeptides
 CC (SPs). The SPs can modulate at least one of proliferation,
 CC differentiation or survival of a cell which expresses the SP. The
 CC products can be used to modulate spermatogenesis, osteogenesis,
 CC chondrogenesis or neuronal cell differentiation or to enhance survival of
 CC neuronal cells, e.g. to prevent apoptosis. They can be used for treating
 CC e.g. degenerative disorders marked by loss of particular cell-types,
 CC apoptosis, neoplastic and/or hyperplastic disorders. In particular they
 CC can be used to treat neurological conditions derived from acute,

CC subacute, or chronic injury to the nervous system, aging of the nervous
CC system chronic neurodegenerative diseases of the nervous system, chronic
CC immunological diseases of the nervous system or affecting the nervous
CC system. They can also treat be used to CNS trauma infarction, infection
CC (such as viral infection with varicella-zoster), metabolic disease,
CC nutritional deficiency, toxic agents (such as cisplatin treatment),
CC chronic pain syndromes or in nerve prostheses for the repair of central
CC and peripheral nerve damage, malignant gliomas, medulloblastomas,
CC neuroectodermal tumours, ependymomas and for liver repair and
CC regeneration of lung tissue in the treatment of emphysema, or for
CC enhancing bond formation, tissue formation or fertility. The products can
CC also be used for detection, diagnosis and drug screening
XX
SQ Sequence 4617 BP; 1338 A; 1003 C; 1008 G; 1268 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 2; Length 4617;
Best Local Similarity 90.0%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCTGAAC 22
| | | | | | | | | | | | | | | | | |
Db 701 AGCTGAACGATAGCTGAAC 720

RESULT 21
ABL18270/c
ID ABL18270 standard; DNA; 5749 BP.

AC ABL18270;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 6283.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 6283; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 5749 BP; 1646 A; 1313 C; 1228 G; 1562 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 4; Length 5749;
Best Local Similarity 90.0%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAGCGGAA 21
| | | | | | | | | | | | | | | | | |
Db 3671 CACCTGAAGTTAAGCGGAA 3652

RESULT 22

ABL05012/c

ID ABL05012 standard; cDNA; 5905 BP.

XX ABL05012;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 9518.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.

OS Drosophila melanogaster.

PN WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX P-PSDB; ABB60909.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 9518; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 5905 BP; 1618 A; 1317 C; 1316 G; 1654 T; 0 U; 0 Other;

Query Match 76.4%; Score 16.8; DB 4; Length 5905;
Best Local Similarity 90.0%; Pred. No. 3.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCGGAA 22
| | | | | | | | | | | | | | | | | |
Db 4259 AGCTGAACGATAGCTGAAC 4240

RESULT 23

ABL03787

ID ABL03787 standard; cDNA; 8114 BP.

XX ABL03787;

XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5843.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KM pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PS 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PR WPI; 2001-656860/75.
XX PS P-PSDB; ABB59684.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 5843; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 8114 BP; 2159 A; 2177 C; 2263 G; 1515 T; 0 U; 0 Other;
XX
XX Query Match 76.4%; Score 16.8; DB 4; Length 8114;
XX Best Local Similarity 90.0%; Pred. No. 3.8e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 3 ACCTGAACGATAGCGGAC 22
XX Db 3740 ACCTGAACGAGAGTGGAC 3759
XX
XX RESULT 24
XX ABL03786/c
XX ID ABL03786 standard; cDNA; 18737 BP.
XX AC ABL03786;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 5840.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KM pharmaceutical; gene; ss.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX

XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PS 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX PR WPI; 2001-656860/75.
XX PS P-PSDB; ABB59683.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 5840; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 18737 BP; 4552 A; 4551 C; 4677 G; 4957 T; 0 U; 0 Other;
XX
XX Query Match 76.4%; Score 16.8; DB 4; Length 18737;
XX Best Local Similarity 90.0%; Pred. No. 4.1e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 3 ACCTGAACGATAGCGGAC 22
XX Db 7842 ACCTGAACGAGAGTGGAC 7823
XX
XX RESULT 25
XX AAL53422
XX ID AAL53422 standard; DNA; 545 BP.
XX AC AAL53422;
XX DT 12-DEC-2002 (first entry)
XX DE Human papillomavirus-18 (HPV18) E6E7-DNA sequence.
XX KW Virucide; cytostatic; E6; E7 fusion protein; HPV; immunogenic; vaccine;
XX KW fusion partner; immunogenicity; HPV infection; neoplasm; HPV18;
XX KW human papillomavirus-18; E6E7-sequence; gene; ds.
XX OS Human papillomavirus.
XX FT Key Location/Qualifiers
XX CDS 14..532
XX FT /*tag= a
XX FT /product= "HPV18 protein"
XX
XX PN EP1243655-A1.
XX PD 25-SEP-2002.
XX PF 23-MAR-2001; 2001EP-00107271.
XX PR 23-MAR-2001; 2001EP-00107271.
XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX PI Cid-Arregui A, Zur Hausen H;
XX

DR WPI; 2002-724952/79.
DR P-PSDB; AAO22924.
XX
PT A new DNA sequence encoding a fusion protein comprising a mutagenized HPV
PT (human papillomavirus) E6 or E7 coding sequence and a sequence encoding a
PT highly immunogenic fusion partner is useful to vaccinate against HPV
PT infection.
XX
XX Disclosure; Fig 4; 34pp; English.
PS
CC The invention relates to a new DNA sequence encodes an E6 or E7 fusion
CC protein of HPV, where at least 20% of the original codons are replaced by
CC codons which lead to enhanced translation in a mammalian cell, containing
CC a mutation which results in production of a truncated non-functional
CC protein, and encoding a highly immunogenic polypeptide fusion partner
CC capable of enhancing immunogenicity of the E6 or E7 protein in the
CC mammalian host. The invention is used as a vaccine for the prevention or
CC treatment of an HPV infection or a neoplasm associated with HPV
CC infection. This polynucleotide sequence represents the DNA coding for the
CC human papillomavirus-18 (HPV18) E6E7-sequence of the invention
XX
XX Sequence 545 BP; 115 A; 214 C; 139 G; 77 T; 0 U; 0 Other;
SQ
Query Match 74.5%; Score 16.4; DB 6; Length 545;
Best Local Similarity 94.4%; Pred. No. 4.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCACCTGAACGATAGCG 18
DB 388 CCACCTGAACGATAGCG 405
RESULT 26
ABL03277/c
ID ABL03277 standard; cDNA; 683 BP.
XX
XX ABL03277;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 4313.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB59174.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 4313; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 683 BP; 204 A; 176 C; 143 G; 160 T; 0 U; 0 Other;
SQ
Query Match 73.6%; Score 16.2; DB 4; Length 683;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CACCTGAACGATAGCGGAAC 22
DB 98 CACTTGAAGATAGCGGAAC 78
RESULT 27
ADK59601/c
ID ADK59601 standard; DNA; 730 BP.
XX
XX ADK59601;
XX
XX 06-MAY-2004 (first entry)
XX
XX Plant DNA sequence which confers altered metabolic characteristic #5984.
XX
XX altered metabolic characteristic; plant; acid metabolism;
XX alcohol metabolism; fatty acid metabolism;
XX branched fatty acid metabolism; alkaloid metabolism;
XX amino acid metabolism; ester metabolism; glyceride metabolism;
XX phenolic metabolism; carbohydrate metabolism; sterol metabolism;
XX terpene metabolism; isoprenoid metabolism; alkene metabolism;
XX alkyne metabolism; hydrocarbon metabolism; ketone metabolism;
XX quinone metabolism; disease resistance; gene shuffling; sexual PCR; ds.
XX
XX Unidentified.
XX
XX WO2003020936-A1.
XX
XX 13-MAR-2003.
XX
XX 30-AUG-2002; 2002WO-US027884.
XX
XX 31-AUG-2001; 2001US-0316471P.
XX
XX (DOWC) DOW CHEM CO.
XX (DOWC) DOW AGROSCIENCES LLC.
XX
XX Weglarz T, Gachotte D, Blakeslee B, McCreary DA, Pell RJ;
XX Oriedo JVB, Crosley R, Reddy AS, Shukla V, Larrinua I, Miller BA;
XX
XX WPI; 2003-313091/30.
XX
XX Novel genes that confer altered metabolic characteristics in Nicotiana
XX benthamiana plants, useful for altering the levels of metabolites e.g.
XX acids, fatty acids, amino acids, carbohydrates, hydrocarbons and sterols.
XX
XX Claim 1; SEQ ID NO 6984; 2576pp; English.
XX
XX The invention comprises DNA sequences which confer an altered metabolic
XX characteristic when they are expressed in a plant. The DNA sequences of
XX the invention are useful for producing plants with an altered metabolic
XX characteristic, such as: altered acid metabolism, alcohol metabolism,
XX fatty acid metabolism, branched fatty acid metabolism, alkaloid or other
XX base metabolism, altered amino acid metabolism, altered ester metabolism,
XX altered glyceride metabolism, altered phenolic metabolism, altered
XX carbohydrate metabolism, altered sterol, oxygenated terpene, or
XX isoprenoid metabolism, alkene or alkyne metabolism, hydrocarbon
XX metabolism, ketone or quinone metabolism. The DNA sequences of the
XX invention may be used to provide disease resistance in a plant and gene
XX shuffling or sexual PCR procedures. The present nucleic acid represents a

```
CC DNA sequence of the invention.
SQ Sequence 730 BP; 175 A; 196 C; 175 G; 184 T; 0 U; 0 Other;

Query Match      73.6%; Score 16.2; DB 10; Length 730;
Best Local Similarity 85.7%; Pred. No. 5.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      2 CACCTGAACGATAGCGGAC 22
      ||||| ||||| |||||
Db      292 CACCTGAACGCTAGCGGAAC 272

RESULT 28
AAC45825/c
ID AAC45825 standard; DNA; 1005 BP.
XX
AC AAC45825;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 47899.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 03-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 23-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
```

```
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.6%; Score 16.2; DB 3; Length 1005;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 73.6%; Score 16.2; DB 3; Length 1005;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
Db 430 CACCTGAACGTAAGCGGAAC 410

RESULT 29
ABZ13014/C
ID ABZ13014 standard; DNA; 1005 BP.
XX
AC ABZ13014;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 819.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
(SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
WPI; 2002-304127/34.
XX
Identifying a stress condition to which a plant cell has been exposed and
producing plants with increased tolerance to these abiotic stresses.
XX
Claim 144; SEQ ID NO 819; 577pp + Sequence Listing; English.
XX
The invention relates to identifying a stress condition to which a plant
cell has been exposed, comprising: (a) contacting nucleic acid
representative of expressed polynucleotides in the plant cell with an
array or probes representative of the plant cell genome; and (b)
detecting a profile of expressed polynucleotides in the plant cell
characteristic of a stress response. The method is useful in the
production of transgenic plants, cells and seeds and in producing plants
with increased tolerance to abiotic stress. The present sequence is that
of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
in methods of the invention. Note: The sequence data for this patent is
not represented in the printed specification but is based on sequence
information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1005 BP; 227 A; 264 C; 248 G; 266 T; 0 U; 0 Other;

Query Match 73.6%; Score 16.2; DB 6; Length 1005;
Best Local Similarity 85.7%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
Db 430 CACCTGAACGTAAGCGGAAC 410

RESULT 30
ADN72926/C
ID ADN72926 standard; cDNA; 1005 BP.
XX
AC ADN72926;
XX
DT 15-JUL-2004 (first entry)
```


PR	24-JUN-1999,	99US-01466959,
PR	28-JUN-1999,	99US-01468239,
PR	29-JUN-1999,	99US-01469919,
PR	30-JUN-1999,	99US-01412879,
PR	01-JUL-1999,	99US-01418429,
PR	01-JUL-1999,	99US-01421549,
PR	02-JUL-1999,	99US-01420559,
PR	06-JUL-1999,	99US-01423909,
PR	08-JUL-1999,	99US-01428039,
PR	09-JUL-1999,	99US-01429209,
PR	13-JUL-1999,	99US-01429779,
PR	13-JUL-1999,	99US-01435429,
PR	14-JUL-1999,	99US-01436249,
PR	15-JUL-1999,	99US-01440059,
PR	16-JUL-1999,	99US-01440859,
PR	16-JUL-1999,	99US-01440869,
PR	19-JUL-1999,	99US-01443259,
PR	19-JUL-1999,	99US-01443319,
PR	19-JUL-1999,	99US-01443329,
PR	19-JUL-1999,	99US-01443339,
PR	19-JUL-1999,	99US-01443349,
PR	19-JUL-1999,	99US-01443359,
PR	20-JUL-1999,	99US-01443529,
PR	20-JUL-1999,	99US-01446329,
PR	20-JUL-1999,	99US-01448849,
PR	21-JUL-1999,	99US-01448149,
PR	21-JUL-1999,	99US-01450869,
PR	21-JUL-1999,	99US-01450889,
PR	22-JUL-1999,	99US-01450859,
PR	22-JUL-1999,	99US-01450879,
PR	22-JUL-1999,	99US-01450899,
PR	22-JUL-1999,	99US-01451929,
PR	23-JUL-1999,	99US-01451459,
PR	23-JUL-1999,	99US-01452189,
PR	23-JUL-1999,	99US-01452249,
PR	26-JUL-1999,	99US-01452769,
PR	27-JUL-1999,	99US-01459139,
PR	27-JUL-1999,	99US-01459189,
PR	27-JUL-1999,	99US-01459199,
PR	28-JUL-1999,	99US-01459519,
PR	02-AUG-1999,	99US-01453869,
PR	02-AUG-1999,	99US-01463889,
PR	02-AUG-1999,	99US-01463899,
PR	03-AUG-1999,	99US-01470389,
PR	04-AUG-1999,	99US-01472049,
PR	04-AUG-1999,	99US-01473029,
PR	05-AUG-1999,	99US-01471929,
PR	06-AUG-1999,	99US-01472609,
PR	06-AUG-1999,	99US-01473039,
PR	06-AUG-1999,	99US-01474169,
PR	09-AUG-1999,	99US-01474939,
PR	09-AUG-1999,	99US-01479359,
PR	10-AUG-1999,	99US-01481719,
PR	11-AUG-1999,	99US-01483199,
PR	12-AUG-1999,	99US-01483419,
PR	13-AUG-1999,	99US-01485569,
PR	13-AUG-1999,	99US-01486849,
PR	16-AUG-1999,	99US-01493689,
PR	17-AUG-1999,	99US-01491759,
PR	18-AUG-1999,	99US-01494269,
PR	20-AUG-1999,	99US-01497229,
PR	20-AUG-1999,	99US-01497239,
PR	20-AUG-1999,	99US-01499299,
PR	23-AUG-1999,	99US-01499029,
PR	23-AUG-1999,	99US-01499309,
PR	25-AUG-1999,	99US-01505669,
PR	26-AUG-1999,	99US-01508849,
PR	27-AUG-1999,	99US-01510659,
PR	27-AUG-1999,	99US-01510669,
PR	30-AUG-1999,	99US-01510309,
PR	31-AUG-1999,	99US-01514389,
PR	01-SEP-1999,	99US-01519309,

PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155113P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157553P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159594P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161993P.
PR	29-OCT-1999;	99US-0162142P.

Query Match	73.6%	Score 16.2;	DB 3;	Length 1053;
Best Local Similarity	85.7%	Pred. No. 6e+02;		
Matches 18: Conservative	0;	Mismatches 3;	Indels 0	

Qy 2 CACCTGAACGATAAGCGGAAC 22
Db 971 CACCTGAACCGTAAGCGGAAC 951

RESULT 32
ABX07448/c
ID ABX07448 standard; DNA; 1101 BP.

AC ABX07448;

DT 27-OCT-2003 (revised)

DI	27-OCT-2003	(revised)
DT	11-FEB-2003	(first entry)

DE S. pneumoniae type 4 strain coding region #1736.

Gene; ds; bacterial meningitis; pneumonia; sepsis; otitis media; ear infection; antiinflammatory; antibacterial; immunostimulant; auditory; respiratory; gene therapy; vaccine.

XX OS Streptococcus pneumoniae: type 4 strain.

XX PN WO200277021-A2

XX PD 03-OCT-2002.
XX PF 27-MAR-2002; 2002WO-1B002163.
XX PR 27-MAR-2001; 2001GB-00007658.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Masignani V, Tettelin H, Fraser C;
XX PF WPI: 2003-040579/03.
XX PR P-PSDB; ABU02159.
XX DR
XX DR
XX PT New proteins and nucleic acid molecules from *Streptococcus pneumoniae*,
XX PT useful as medicaments for treating or preventing a disease or infection
XX PT due to *Streptococcus* bacteria, such as pneumonia, sepsis, otitis media or
XX PT ear infection.
XX PS Claim 6; SEQ ID NO 3471; 56pp; English.
XX XX
XX CC The invention relates to a protein comprising or having at least 50%
XX CC identity to any of the 2469 amino acid sequences, identified in the
XX CC specification (available on a computer readable format), or its fragment,
XX CC expressed from 2469 of 2489 identified DNA coding regions from the
XX CC *Streptococcus pneumoniae* type 4 strain genomic sequence appearing as
XX CC ABS56454. Also included are an antibody which binds one of the proteins,
XX CC treating a patient by administering the protein, DNA or antibody (in a
XX CC composition), a kit comprising first and second primers, which are the
XX CC nucleic acid cited above or fragments between nucleotides 8-100 of a
XX CC sequence not defined in the specification, for amplifying a target
XX CC sequence contained within a *Streptococcus* nucleic acid sequence, where
XX CC the first primer is substantially complementary to the target sequence
XX CC and the second primer is substantially complementary to the complement of
XX CC the target sequence, and where the parts of the primers having
XX CC substantial complementarity define the termini of the target sequence to
XX CC be amplified, assay comprising contacting a test compound with the
XX CC protein, and determining whether the test compound binds to the protein
XX CC and a *Streptococcus pneumoniae* bacterium, where one or more genes
XX CC encoding the proteins has been rendered inactive. The proteins, nucleic
XX CC acid molecules, antibody and compositions are useful as medicaments for
XX CC treating or preventing a disease or infection due to *Streptococcus*
XX CC bacteria, particularly *S. pneumoniae*, such as pneumonia, sepsis, otitis
XX CC media or ear infection. They are also useful in developing vaccines,
XX CC diagnostics and antibiotics. The methods are useful for identifying
XX CC immunodominant proteins. The present sequence is one of the 2489
XX CC identified coding region from the genomic sequence. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 27-OCT-2003 to
XX CC standardise OS field)
XX SQ Sequence 1101 BP; 349 A; 175 C; 259 G; 318 T; 0 U; 0 Other;

Query Match 73.6%; Score 16.2; DB 10; Length 1101;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGACGATAGCGGAC 22
DB 709 CACCCCAACGATAGCTGAAC 689

RESULT 33
ACA50068/C
ID ACA50068 standard; DNA; 1104 BP.
XX AC
XX AC ACA50068;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #31725.

XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS
XX OS Streptococcus pneumoniae.
XX PN WO200277183-A2.
XX XX
XX PD 03-OCT-2002.
XX XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX XX
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX XX
XX PA (ELIT-) ELITRA PHARM INC.
XX XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI: 2003-029926/02.
XX DR P-PSDB; ABU46198.
XX XX
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 37938; 1766pp; English.
XX XX
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1104 BP; 351 A; 177 C; 258 G; 318 T; 0 U; 0 Other;

Query Match 73.6%; Score 16.2; DB 8; Length 1104;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGACGATAGCGGAC 22
DB 709 CACCCCAACGATAGCTGAAC 689

RESULT 34
ADR93541/c
ID ADR93541 standard; DNA; 1128 BP.
XX AC ADR93541;
XX DT 16-DEC-2004 (first entry)
XX DE Novel S. pneumoniae DNA sequence, SEQ ID 2176.
XX KW Meningitis; bacteraemia; pneumonia; otitis media; ds;
XX KW bacterial infection.
XX OS Streptococcus pneumoniae.
XX PN US6800744-B1.
XX PD 05-OCT-2004.
XX PF 30-JUN-1998; 98US-00107433.
XX PR 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX PA Doucette-Stamm LA, Bush D;
PI WPI; 2004-697205/69.
DR P-PSDB; ADR96144.
XX New isolated nucleic acid encoding a Streptococcus pneumoniae
PT polypeptide, useful for diagnosing, preventing and/or treating
PT pathological conditions resulting from the bacterial infection.
XX Disclosure; SEQ ID NO 2176; 151pp; English.
XX The invention relates to an isolated nucleic acid comprising a sequence
CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR93366, ADR92650 or
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
CC hybridisable under high stringency conditions to the nucleotide sequence.
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
CC Also included are a recombinant expression vector comprising the isolated
CC nucleic acid cited above operably linked to a transcription regulatory
CC element, a cell comprising the recombinant expression vector and a probe
CC comprising at least 20 consecutive nucleotides of the nucleotide
CC sequences as cited above. The methods and compositions of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC pathological conditions resulting from bacterial infection by
CC Streptococcus pneumoniae e.g. pneumonia, bacteraemia, meningitis and
CC otitis media. The present sequence is one of the 2603 disclosed S.
CC pneumoniae nucleic acid sequences. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=6800744B1.
XX SQ Sequence 1128 BP; 363 A; 178 C; 263 G; 324 T; 0 U; 0 Other;
Query Match 73.6%; Score 16.2; DB 13; Length 1128;
Best Local Similarity 85.7%; Pred. No. 6e-02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 CACCTGAACGATAGCGGAC 22
Db 733 CACCCCAACGATAGCTGAAC 713

RESULT 35
ADA15526/c
ID ADA15526 standard; DNA; 1699 BP.
XX AC ADA15526;
XX DT 06-NOV-2003 (first entry)
XX DE DNA encoding A. thaliana protein, SEQ ID NO:81.
XX KW Transgenic plant; plant; transcription factor; trait; root;
XX KW plant stability; ds; gene.
XX OS Arabidopsis thaliana.
XX PN US2003061637-A1.
XX PD 27-MAR-2003.
XX PF 21-OCT-2002; 2002US-00278173.
XX PR 23-MAR-1999; 99US-0125814P.
XX PR 22-MAR-2000; 2000US-00533393.
XX (JIAN/) JIANG C.
XX PA (BROU/) BROUN P.
XX PA (RIEC/) RIECHMANN J L.
XX PA (PINE/) PINEDA O.
XX PA (ZHAN/) ZHANG J.
XX PA (YUGG/) YU G.
XX PA (PIILG/) PILGRIM M.
XX PA (KEDD/) KEDDIE J.
XX PA (HEAR/) HEARD J.
XX PA (REUB/) REUBER L.
XX PA (RATC/) RATCLIFFE O.
XX PA (ADAM/) ADAM L.
XX PA (SAMA/) SAMAHA R.
XX PI Jiang C, Broun P, Riechmann J, Pineda O, Zhang J, Yu G;
PI Pilgrim M, Keddie J, Heard J, Reuber L, Ratcliffe O, Adam L;
PI Samaha R;
XX WPI; 2003-555503/52.
XX P-PSDB; ADA15527.
XX New transgenic plant comprising a recombinant polynucleotide, useful for
PT altering a plant's trait for increasing plant stability.
XX Claim 14; SEQ ID NO 81; 159pp; English.
XX The invention discloses a transgenic plant having an A. thaliana
CC recombinant polynucleotide (a transcription factor) comprising a sequence
CC encoding a polypeptide with at least 6 consecutive amino acids of one of
CC the amino acid sequences given in the specification. The recombinant
CC polynucleotide alters a trait of the transgenic plant's roots when
CC compared to the same trait of the roots of another plant lacking the
CC recombinant polynucleotide. Also claimed are methods for altering the
CC expression levels of at least one gene of a plant, altering a trait
CC associated with a plant's roots and altering a plant's trait. The
CC recombinant polynucleotide further comprises a promoter operably linked
CC to the nucleotide sequence. The promoter is constitutive or inducible or
CC root-active. The method for altering a trait associated with roots
CC comprises transforming a plant with the recombinant polynucleotide,
CC selecting the transformed plants and identifying a transformed plant with
CC roots having an altered trait. The method for altering the expression
CC levels of at least one gene of a plant comprises transforming a plant
CC with the recombinant polynucleotide and selecting the transformed plant.
CC The method for altering a trait associated with a plant's roots comprises
CC transforming a plant with the recombinant polynucleotide and selecting
CC the transformed plant. The method for altering a plant's trait comprises
CC providing a database sequence comparing the database sequence with the
CC polypeptide, selecting the database that meets selected sequence criteria

CC and transforming the selected database sequence in the plant. The methods
 CC also comprise providing a test polynucleotide, hybridising the test
 CC polynucleotide with the recombinant polynucleotide at low stringency and
 CC transforming the hybridising test polynucleotide in a plant to alter a
 CC trait of the plant. The transgenic plant is useful for altering a plant's
 CC trait for increasing plant stability. The sequence presented is a DNA
 CC encoding a protein of the invention.

XX SQ Sequence 1699 BP; 396 A; 364 C; 357 G; 582 T; 0 U; 0 Other;

Query Match 73.6%; Score 16.2; DB 9; Length 1699;

Best Local Similarity 85.7%; Pred. No. 6.3e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAGCGGAAC 22

||||||| |||||||
 DB 846 CACCTGAACGATAGCGGAAC 826

RESULT 36

ADB31770/c

ID ADB31770 standard; DNA; 1699 BP.

XX AC ADB31770;

XX DT 04-DEC-2003 (first entry)

XX DE DNA encoding plant (A. thaliana) transcription factor polypeptide #2.

XX KW Plant; transcription factor; transgenic plant; transgenic; plant trait;

XX KM modified trait; gene; ds.

XX OS Arabidopsis thaliana.

XX PN US2003101481-A1.

XX PD 29-MAY-2003.

XX PF 15-NOV-2002; 2002US-00295403.

XX PR 22-SEP-1998; 98US-0101349P.

XX PR 06-OCT-1998; 98US-0103312P.

XX PR 17-NOV-1998; 98US-0108734P.

XX PR 22-DEC-1998; 98US-0113409P.

XX PR 13-SEP-1999; 99US-00394519.

XX PA (ZHAN/) ZHANG J.

XX PA (FROM/) FROMM M.

XX PA (HEAR/) HEARD J.

XX PA (RIEC/) RIECHMANN J L.

XX PA (ADAM/) ADAM L.

XX PA (BROU/) BROUN P.

XX PA (PINE/) PINEDA O.

XX PA (REUB/) REUBER L.

XX PA (KEDD/) KEDDIE J.

XX PA (YUGG/) YU G.

XX PA (JIAN/) JIANG C.

XX PI Zhang J, Fromm M, Heard J, Riechmann JL, Adam L, Broun P;

XX PI Pineda O, Reuber L, Keddie J, Yu G, Jiang C;

XX XX

DR WPI; 2003-597572/56.

DR P-PSDB; ADB31771.

XX XX

PT New isolated polypeptides and polynucleotide sequences, useful for

PT screening a molecule to identify a molecule that modifies plant trait,

PT and for producing plants with modified traits.

XX XX

PS Claim 1; SEQ ID NO 3; 17pp; English.

XX XX

CC The present invention relates to the isolation of plant (Arabidopsis

CC thaliana) transcription factor polypeptide and polynucleotide sequences.

CC Also disclosed are: an expression vector comprising the isolated

CC polynucleotide, a host cell comprising the expression vector, a
 CC transgenic plant comprising the isolated polynucleotide, a transgenic
 CC plant ectopically expressing the isolated polynucleotide or polypeptide,
 CC a method for screening a molecule to identify a molecule that modifies a
 CC plant trait by placing the molecule in contact with the plant, and
 CC monitoring the effect of the molecule on the expressing or activity of
 CC the polypeptide or polynucleotide, and producing a transgenic plant
 CC having a modified trait by ectopically expressing the isolated
 CC polypeptide and selecting a plant with the modified trait. The
 CC polypeptides, polynucleotides and methods are useful for screening a
 CC molecule to identify a molecule that modifies plant trait, and for
 CC producing plants with modified traits. The present sequence represents a
 CC plant transcription factor polynucleotide of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site at seqdata.uspto.gov.

XX SQ Sequence 1699 BP; 396 A; 364 C; 357 G; 582 T; 0 U; 0 Other;

Query Match 73.6%; Score 16.2; DB 10; Length 1699;

Best Local Similarity 85.7%; Pred. No. 6.3e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAGCGGAAC 22

||||||| |||||||
 DB 846 CACCTGAACGATAGCGGAAC 826

RESULT 37

ADD30689/c

ID ADD30689 standard; cDNA; 1699 BP.

XX AC ADD30689;

XX DT 15-JAN-2004 (first entry)

XX DE Plant yield-related polynucleotide clone G5.

XX KW ds; transcription factor; transgenic plant; growth rate; senescence;

XX KW seed germination rate; plant vigor; seedling vigor.

XX OS Arabidopsis thaliana.

XX PN WO2003013227-A2.

XX PD 20-FEB-2003.

XX PF 09-AUG-2002; 2002WO-US025805.

XX PR 09-AUG-2001; 2001US-0310847P.

XX PR 19-NOV-2001; 2001US-0336049P.

XX PR 11-DEC-2001; 2001US-0338692P.

XX PR 14-JUN-2002; 2002US-00171468.

XX PA (MEND-) MENDEL BIOTECHNOLOGY INC.

XX XX

PI Ratcliffe O, Riechmann JL, Adam LJ, Dubell AT, Heard JE;

PI Pilgrim ML, Jiang C, Reuber TL, Creelman RA, Pineda O, Yu G;

PI Broun PB;

XX XX

DR WPI; 2003-248221/24.

DR P-PSDB; ADD30690.

XX XX

PT New plant transcription factor polynucleotides and polypeptides, useful

PT in producing transgenic plants with commercially valuable properties,

PT such as an alteration in a plant growth characteristic, e.g. growth rate

PT or apomixis.

XX XX

PS Disclosure; SEQ ID NO 718; 454pp; English.

XX XX

CC The invention relates to a number of isolated Arabidopsis thaliana cDNA

CC sequences and their encoded proteins which are especially transcription

CC factor related cDNA's and proteins. The isolated or recombinant plant

CC transcription factor polynucleotides and polypeptides are useful in
 CC producing transgenic plants with commercially valuable properties, i.e.
 CC modified or altered desirable traits as compared to a reference plant,
 CC such as an alteration in a plant growth characteristic, e.g. growth rate,
 CC germination rate of seeds, vigor of plants and seedlings, or leaf and
 CC flower senescence. Sequence information related to the polynucleotides
 CC and polypeptides can also be used in bioinformatic search methods. The
 CC transgenic plant is useful for growing a progeny plant from a parent
 CC plant. This sequence represents one of the cDNAs of the invention.
 XX
 SQ Sequence 1699 BP; 396 A; 364 C; 357 G; 582 T; 0 U; 0 Other;
 Query Match 73.6%; Score 16.2; DB 10; Length 1699;
 Best Local Similarity 85.7%; Pred. No. 6.3e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 CACCTGAACGATTAAGCGGAAC 22
 ||||| ||||| ||||| ||||| |||||
 Db 846 CACCTGAACGTAAGCGGAAC 826
 RESULT 38
 ADI43484/C
 ID ADI43484 standard; DNA; 1699 BP.
 XX
 AC ADI43484;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Plant transcription factor polynucleotide #1333.
 XX
 KW transgenic; plant; enhanced tolerance to abiotic stress;
 KW glyphosate tolerance; hormone sensitivity; disease resistance;
 KW sugar sensing; flowering; flower structure; stem bifurcation;
 KW branching pattern; apical dominance; trichome; stem morphology;
 KW root growth; root hair; seed development; cell proliferation;
 KW cell differentiation; premature senescence; necrosis; plant size;
 KW leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
 KW plant anthocyanin; light response; shade avoidance; bioinformatic;
 KW transcription factor; gene; ds.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2004019927-A1.
 XX
 PD 29-JAN-2004.
 XX
 PF 25-FEB-2003; 2003US-00374780.
 XX
 PR 18-APR-2001; 2001US-00837944.
 XX
 PA (SHER/) SHERMAN B K.
 PA (RIEC/) RIECHMANN J L.
 PA (JIAN/) JIANG C.
 PA (HEAR/) HEARD J E.
 PA (HAAR/) HAAKE V.
 PA (CREE/) CREELMAN R A.
 PA (RATC/) RATCLIFFE O.
 PA (ADAM/) ADAM L J.
 PA (REUB/) REUBER T L.
 PA (KEDD/) KEDDIE J.
 PA (BROU/) BROUN P E.
 PA (PILG/) PILGRIM M L.
 PA (DUBE/) DUBELL A N.
 PA (FINE/) PINEDA O.
 PA (YUGG/) YU G.
 XX
 PI Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
 PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddie J, Broun PE;
 PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
 XX
 XX WPI; 2004-132245/13.
 DR
 XX

PT New transgenic plant comprising a recombinant polynucleotide of any one
 PT of more than 500 nucleotide sequences, useful in bioinformatic search
 PT methods.
 PS
 XX Claim 1; SEQ ID NO 1947; 435pp; English.
 XX
 CC The invention describes a transgenic plant comprising a recombinant
 CC polynucleotide of any one of more than 500 nucleotide sequences fully
 CC defined in the specification or its complement. The method of the
 CC invention can be used to produced a plant having altered traits such as:
 CC enhanced tolerance to abiotic stress; glyophosphate tolerance; hormone
 CC sensitivity; disease resistance; sugar sensing; early or late flowering;
 CC altered flower structure, change in stem bifurcations, altered branching
 CC pattern, reduced apical dominance, reduced trichome density; lack of
 CC trichomes; reduced ectopic trichome development; altered trichome
 CC development; increase in trichome number; altered stem morphology;
 CC increased root growth; increased root hairs; altered seed development;
 CC altered cell proliferation or cell differentiation; rapid development;
 CC premature senescence; increased necrosis; increase in seedling or plant
 CC size; decreased plant size; leaf morphology; seed morphology; seed
 CC biochemistry; increase in root anthocyanins; increase in plant
 CC anthocyanins, or alteration in light response or shade avoidance. The
 CC transgenic plant, polynucleotides and polypeptides are useful in
 CC bioinformatic search methods. This sequence represents a plant
 CC transcription factor, and an orthologue of Arabidopsis thaliana
 CC transcription factors isolated in the invention, that can be used in the
 CC creation of a transgenic plant with altered traits.
 XX
 SQ Sequence 1699 BP; 396 A; 364 C; 357 G; 582 T; 0 U; 0 Other;
 Query Match 73.6%; Score 16.2; DB 12; Length 1699;
 Best Local Similarity 85.7%; Pred. No. 6.3e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 CACCTGAACGATTAAGCGGAAC 22
 ||||| ||||| ||||| ||||| |||||
 Db 846 CACCTGAACGTAAGCGGAAC 826
 RESULT 39
 ADO03298/C
 ID ADO03298 standard; cDNA; 1699 BP.
 XX
 AC ADO03298;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Thalecress transcription factor cDNA #481.
 XX
 KW Thalecress; transcription factor; ss; gene; plant; transgenic;
 KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
 KW phosphate limitation; potassium limitation; nitrogen limitation;
 KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
 KW flowering; inflorescence architectural change;
 KW meristem cell differentiation; phyllotaxy; apical dominance;
 KW trichome development; seed development; premature senescence;
 KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
 KW seed morphology; secondary metabolism; light response; shade avoidance.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2004045049-A1.
 XX
 PD 04-MAR-2004.
 XX
 PF 10-APR-2003; 2003US-00412699.
 XX
 PR 13-SEP-1999; 99US-00394519.
 PR 21-JAN-2000; 2000US-00489376.
 PR 17-FEB-2000; 2000US-00506720.
 PR 22-MAR-2000; 2000US-00532591.
 PR 22-MAR-2000; 2000US-00533029.
 PR 22-MAR-2000; 2000US-00533030.

PR 22-MAR-2000; 2000US-00533392.
 PR 22-MAR-2000; 2000US-005333648.
 PR 06-APR-2000; 2000WO-US009448.
 PR 16-NOV-2000; 2000US-00713994.
 PR 27-MAR-2001; 2001US-00819142.
 PR 17-APR-2001; 2001US-00837444.
 PR 30-JAN-2002; 2002US-00958131.
 PR 14-JUN-2002; 2002US-00171468.
 PR 09-AUG-2002; 2002US-00225066.
 PR 09-AUG-2002; 2002US-00225067.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 25-FEB-2003; 2003US-00374780.
 XX
 (ZHAN/) ZHANG J.
 (FROM/) FROMM M E.
 (HEAR/) HEARD J E.
 (RIEC/) RIECHMANN J L.
 (ADAM/) ADAM L J.
 (BROU/) BROUN P E.
 (PINE/) PINEDA O.
 (REUB/) REUBER T L.
 (KEDD/) KEDDIE J S.
 (YUGG/) YU G.
 (JIAN/) JIANG C.
 (SAMA/) SAMAHA R S.
 (PILG/) PILGRIM M L.
 (CREE/) CREELMAN R A.
 (DUBE/) DUBELL A N.
 (RATC/) RATCLIFFE O.
 (KUMI/) KUMIMOTO R.
 (SHER/) SHERMAN B K.
 XX
 Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE,
 PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
 PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
 PI Sherman BK;
 XX WPI: 2004-225755/21.
 DR P-PSDB; ADO03299.
 XX
 PT New transgenic plant, useful in developing phenotypes with altered or
 PT improved characteristics or traits.
 XX
 PS Claim 1; SEQ ID NO 1712; 213pp; English.
 XX
 CC The invention relates to a transgenic plant comprising a recombinant
 CC polynucleotide having a polynucleotide sequence or its complementary
 CC sequence comprising a sequence encoding a polypeptide, that initiates
 CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
 CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01588
 CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic
 CC plant to grow a progeny plant, an expression cassette (comprising a
 CC constitutive, inducible or tissue-specific promoter and a recombinant
 CC polynucleotide described above), a host cell comprising the expression
 CC cassette, producing a modified plant having a modified trait, identifying
 CC a factor that is modulated by or interacts with a polypeptide encoded by
 CC the polynucleotide sequence and identifying at least one downstream
 CC polynucleotide sequence that is subject to a regulatory effect of any of
 CC the polypeptides encoded by the polynucleotide described above. The
 CC transgenic plant is useful for producing a plant that has an altered
 CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
 CC to chilling, germination in cold conditions, freezing tolerance, tolerance
 CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
 CC salt, tolerance to phosphate limitation, tolerance to potassium
 CC limitation, decreased sensitivity to nitrogen limitation), altered
 CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
 CC response to ethylene, disease resistance, altered susceptibility to
 CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
 CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
 CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
 CC germination and seedling vigor, early flowering, late flowering, extended
 CC period of flowering, an inflorescence architectural change, a change in

CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
 CC differentiation, altered phyllotaxy, altered branching pattern, reduced
 CC apical dominance, reduced trichome density, ectopic trichome development,
 CC altered trichome development, altered stem morphology, increased root
 CC growth, increased root hairs, altered seed development, altered cell
 CC proliferation/cell differentiation, premature senescence, delayed
 CC senescence, lethality, increased necrosis, an increase in seedling or
 CC plant size, decreased plant size, a change in leaf morphology, increased
 CC altered leaf development, increased leaf size and mass, glossy leaves,
 CC leaf cell expansion, change in seed morphology, altered seed coloration,
 CC increased seed size, decreased seed size, altered seed shape, change in
 CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
 CC content, increased leaf insoluble sugars, decreased leaf insoluble
 CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
 CC content, an alteration of leaf glucosinolate content, change in seed
 CC biochemistry, an increase in seed oil content, decrease in seed fatty acid
 CC content, increase in seed fatty acid content, decrease in seed fatty acid
 CC content, increase in seed protein content, decrease in seed protein
 CC content, alteration in seed prenyl lipid content, increase in seed
 CC sterols, upregulation of genes involved in secondary metabolism, increase
 CC in root anthocyanins, increase in plant anthocyanins, and alteration in
 CC light response or shade avoidance. The present sequence encodes a
 CC thalacress transcription factor of the invention.
 XX
 SQ Sequence 1699 BP; 396 A; 364 C; 357 G; 582 T; 0 U; 0 Other;
 Query Match 73.6%; Score 16.2; DB 12; Length 1699;
 Best Local Similarity 85.7%; Pred. No. 6.3e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 CACCTGAACGATTAAGCGAAC 22
 DB 846 CACCTGAACCGTAAGCGAAC 826
 RESULT 40
 ADO01592/C
 ID ADO01592 standard; cDNA; 1699 BP.
 XX
 AC ADO01592;
 XX
 DT 01-JUL-2004 (first entry)
 XX
 DE Thalacress transcription factor cDNA #3.
 XX
 KW Thalacress; transcription factor; ss; gene; plant; transgenic;
 KW abiotic stress; cold tolerance; heat tolerance; drought; osmotic stress;
 KW phosphate limitation; potassium limitation; nitrogen limitation;
 KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
 KW flowering; inflorescence architectural change;
 KW meristem cell differentiation; phyllotaxy; apical dominance;
 KW trichome development; seed development; premature senescence;
 KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
 KW seed morphology; secondary metabolism; light response; shade avoidance.
 XX Arabidopsis thaliana.
 OS
 XX
 XX US2004045049-A1.
 PN
 XX
 XX 04-MAR-2004.
 PD
 XX
 XX 10-APR-2003; 2003US-00412699.
 PF
 XX
 XX 13-SEP-1999; 99US-00394519.
 PR 21-JAN-2000; 2000US-00489376.
 PR 17-FEB-2000; 2000US-00506720.
 PR 22-MAR-2000; 2000US-00532591.
 PR 22-MAR-2000; 2000US-00533029.
 PR 22-MAR-2000; 2000US-00533030.
 PR 22-MAR-2000; 2000US-00533392.
 PR 22-MAR-2000; 2000US-00533648.
 PR 06-APR-2000; 2000WO-US009448.
 PR 16-NOV-2000; 2000US-00713994.

PR 27-MAR-2001; 2001US-00819142.
PR 17-APR-2001; 2001US-00837444.
PR 30-JAN-2002; 2002US-00958131.
PR 14-JUN-2002; 2002US-00171468.
PR 09-AUG-2002; 2002US-00225066.
PR 09-AUG-2002; 2002US-00225067.
PR 09-AUG-2002; 2002US-00225068.
PR 17-DEC-2002; 2002US-0434166P.
PR 25-FEB-2003; 2003US-00374780.

XX (ZHAN/) ZHANG J.
PA (FROM/) FROMM M E.
PA (HEAR/) HEARD J E.
PA (RIEC/) RIECHMANN J L.
PA (ADAM/) ADAM L J.
PA (BROU/) BROUN P E.
PA (PINE/) PINEDA O.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J S.
PA (YUGG/) YU G.
PA (JIAN/) JIANG C.
PA (SAMA/) SAMAHA R S.
PA (PILG/) PILGRIM M L.
PA (CREE/) CREELMAN R A.
PA (DUBE/) DUBELL A N.
PA (RATC/) RATCLIFFE O.
PA (KUMI/) KUMIMOTO R.
PA (SHER/) SHERMAN B K.

XX Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
PI Sherman BK;

XX WPI: 2004-225755/21.
DR P-PSDB; ADO01593.

PT New transgenic plant, useful in developing phenotypes with altered or
PT improved characteristics or traits.

XX Claim 1; SEQ ID NO 5; 213pp; English.

XX The invention relates to a transgenic plant comprises a recombinant
CC polynucleotide having a polynucleotide sequence or its complementary
CC sequence comprising a sequence encoding a polypeptide, that initiates
CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
CC Rice, Rape or Corn, comprising any of the sequences appearing as ADO01598
CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic
CC plant to grow a progeny plant, an expression cassette (comprising a
CC constitutive, inducible or tissue-specific promoter and a recombinant
CC polynucleotide described above), a host cell comprising the expression
CC cassette, producing a modified plant having a modified trait, identifying
CC a factor that is modulated by or interacts with a polypeptide encoded by
CC the polynucleotide sequence and identifying at least one downstream
CC polynucleotide sequence that is subject to a regulatory effect of any of
CC the polypeptides encoded by the polynucleotide described above. The
CC transgenic plant is useful for producing a plant that has an altered
CC trait e.g. an enhanced tolerance to abiotic stress (increased tolerance
CC to chilling, germination in cold conditions, freezing tolerance, tolerance
CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
CC salt, tolerance to phosphate limitation, tolerance to potassium
CC limitation, decreased sensitivity to nitrogen limitation), altered
CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
CC response to ethylene, disease resistance, altered susceptibility to
CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
CC germination and seedling vigor, early flowering, late flowering, extended
CC period of flowering, an inflorescence architectural change, a change in
CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
CC differentiation, altered phyllotaxy, altered branching pattern, reduced
CC apical dominance, reduced trichome density, ectopic trichome development,
CC altered trichome development, altered stem morphology, increased root

CC growth, increased root hairs, altered seed development, altered cell
CC proliferation/cell differentiation, premature senescence, delayed
CC senescence, lethality, increased necrosis, an increase in seedling or
CC plant size, decreased plant size, a change in leaf morphology, increased
CC altered leaf development, increased leaf size and mass, glossy leaves,
CC leaf cell expansion, change in seed morphology, altered seed coloration,
CC increased seed size, decreased seed size, altered seed shape, change in
CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
CC content, increased leaf insoluble sugars, decreased leaf insoluble
CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid
CC content, an alteration of leaf glucosinolate content, change in seed
CC biochemistry, an increase in seed oil content, decrease in seed oil
CC content, increase in seed fatty acid content, decrease in seed fatty acid
CC content, increase in seed protein content, decrease in seed protein
CC content, alteration in seed prenyl lipid content, increase in seed
CC sterols, upregulation of genes involved in secondary metabolism, increase
CC in root anthocyanins, increase in plant anthocyanins, and alteration in
CC light response or shade avoidance. The present sequence encodes a
CC thalacress transcription factor of the invention.

XX
SQ Sequence 1699 BP; 396 A; 364 C; 357 G; 582 T; 0 U; 0 Other;

Query Match 73.6%; Score 16.2; DB 12; Length 1699;
Best Local Similarity 85.7%; Pred. NO. 6.3e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATTAAGCGGAAC 22

DB 846 CACCTGAACGATTAAGCGGAAC 826

Search completed: June 4, 2005, 07:28:51
Job time : 169.757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:48:59 ; Search time 48.9467 Seconds
(without alignments)
735.454 Million cell updates/sec

Title: US-09-674-277-21
Perfect score: 22
Sequence: 1 ccactgaacgataagcggaac 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18.8	85.5	4450	3	US-08-617-860B-2
C 2	16.8	76.4	1840	2	US-08-484-993B-15
C 3	16.8	76.4	1840	2	US-08-484-158B-15
C 4	16.8	76.4	1840	2	US-08-484-596A-15
C 5	16.8	76.4	1840	2	US-08-480-150A-15
C 6	16.8	76.4	1840	3	US-08-458-731-15
C 7	16.8	76.4	1840	3	US-08-149-223A-15
C 8	16.2	73.6	461	4	US-09-621-976-18882
C 9	16.2	73.6	1070	4	US-09-270-767-13955
C 10	16.2	73.6	1104	4	US-09-583-110-515
C 11	16.2	73.6	1128	4	US-09-107-433-2176
C 12	16.2	73.6	2664	1	US-08-215-709-2
C 13	16.2	73.6	26385	3	US-08-961-527-3
C 14	16.2	73.6	51967	4	US-09-949-016-16982
C 15	15.8	71.8	1020	4	US-10-061-943A-1
C 16	15.8	71.8	1866	4	US-09-799-451-459
C 17	15.6	70.9	53	2	US-08-661-052-8
C 18	15.6	70.9	53	3	US-09-188-082-8
C 19	15.6	70.9	53	3	US-09-364-088-8
C 20	15.6	70.9	53	3	US-09-102-716-8
C 21	15.6	70.9	585	4	US-09-489-039A-1339
C 22	15.6	70.9	594	4	US-09-489-039A-95
C 23	15.6	70.9	705	3	US-08-998-416-821
C 24	15.6	70.9	906	4	US-09-489-039A-101
C 25	15.6	70.9	1290	4	US-09-489-039A-6630
C 26	15.6	70.9	1347	4	US-09-489-039A-6565
C 27	15.6	70.9	1350	4	US-09-252-991A-6609

C 28	15.6	70.9	1353	4	US-09-252-991A-1684	Sequence 1684, Ap
C 29	15.6	70.9	1398	4	US-09-252-991A-1432	Sequence 1432, Ap
C 30	15.6	70.9	1482	4	US-09-252-991A-6692	Sequence 6692, Ap
C 31	15.6	70.9	1791	4	US-09-252-991A-6662	Sequence 6662, Ap
C 32	15.6	70.9	2127	4	US-09-902-540-3949	Sequence 3949, Ap
C 33	15.6	70.9	21295	4	US-09-902-540-1194	Sequence 1194, Ap
C 34	15.6	70.9	80355	4	US-09-949-016-12735	Sequence 12735, A
C 35	15.6	70.9	80357	4	US-09-949-016-13572	Sequence 13572, A
C 36	15.6	70.9	96109	4	US-09-596-002-35	Sequence 35, Appl
C 37	15.4	70.0	765	4	US-09-540-236-1876	Sequence 1876, Ap
C 38	15.4	70.0	3945	4	US-09-602-777A-359	Sequence 359, App
C 39	15.2	69.1	269223	4	US-09-596-002-41	Sequence 41, Appl
C 40	15.2	69.1	662	4	US-09-270-767-1176	Sequence 1176, Ap
C 41	15.2	69.1	682	4	US-09-270-767-16458	Sequence 16458, A
C 42	15.2	69.1	765	4	US-09-134-000C-794	Sequence 794, App
C 43	15.2	69.1	1228	3	US-09-071-035-315	Sequence 315, App
C 44	15.2	69.1	1286	4	US-09-902-540-6514	Sequence 6514, Ap
C 45	15.2	69.1	1326	4	US-09-134-000C-3096	Sequence 3096, Ap

ALIGNMENTS

RESULT 1
US-08-617-860B-2/c
; Sequence 2, Application US/08617860B
; Patent No. 6133506
; GENERAL INFORMATION:
; APPLICANT: Typfer, R., Bautor, J., Bothmann, H., Filsak, E.,
; APPLICANT: Hrvicke-Grandpierre, C., Klein, B., Martini, N.,
; APPLICANT: Mller, A., Schulte, W., Voetz, M., Walek, J.,
; APPLICANT: Schell, J.
; TITLE OF INVENTION: Promoters
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steinberg, Raskin & Davidson, P.C.
; STREET: 1140 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,860B
; FILING DATE: 01-MAR-1996
; PRIOR APPLICATION DATA: PCT/EP94/02950
; APPLICATION NUMBER: 05-SEP-1994
; FILING DATE: 05-SEP-1994
; APPLICATION NUMBER: DE P4329951.2
; FILING DATE: 04-SEP-1993
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4450 base pairs
; TYPE: Nucleic acid
; STRANDEDNESS: double stranded
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: Brassica napus
; IMMEDIATE SOURCE:
; LIBRARY: genomic Lambda FIX II
; CLONE: BnAccasegi
; FEATURE:
; NAME/KEY: CAAT-Signal
; LOCATION: 3124..3127
; FEATURE:
; NAME/KEY: TATA-Signal

/ LOCATION: 3328..3333
/ FEATURE:
/ NAME/KEY: Transcription start
/ LOCATION: 3367
/ FEATURE:
/ NAME/KEY: Startcodon
/ LOCATION: 4089..4091
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 4089..4421
US-08-617-860B-2

Query Match 85.5%; Score 18.8; DB 3; Length 4450;
Best Local Similarity 90.9%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCACCTGACGATAAGCGGAAC 22
DB 4425 CCACCTCAGCGATAAGCTGAAC 4404

RESULT 2
US-08-484-993B-15/c
; Sequence 15, Application US/08484993B
; Patent No. 5837497
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,993B
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Felis domesticus
; DEVELOPMENTAL STAGE: Juvenile

/ HAPLOTYPE: Diploidy
/ TISSUE TYPE: Ovary
/ CELL TYPE: Oocyte
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 57..1766
US-08-484-993B-15

Query Match 76.4%; Score 16.8; DB 2; Length 1840;
Best Local Similarity 90.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGACGATAAGCGGAA 21
DB 1061 CACCTGACGCTAAGCGGGA 1042

RESULT 3
US-08-484-158B-15/c
; Sequence 15, Application US/08484158B
; Patent No. 5976545
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Pharmaceutical Compositions for
; IMMUNOCONTRACEPTION
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,158B
; FILING DATE: 07-JUNE-95
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 09-NOV-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 32794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Felis domesticus

```
; DEVELOPMENTAL STAGE: Juvenile
; HAPLOTYPE: Diploidy
; TISSUE TYPE: Ovary
; CELL TYPE: Oocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1766
US-08-484-158B-15

Query Match 76.4%; Score 16.8; DB 2; Length 1840;
Best Local Similarity 90.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAA 21
|||||
Db 1061 CACCTGAACGCTAAGCGGGA 1042

RESULT 4
US-08-484-596A-15/c
; Sequence 15, Application US/08484596A
; Patent No. 5981228
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,596A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/149,223
; FILING DATE: 11-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Felis domesticus
; DEVELOPMENTAL STAGE: Juvenile
; HAPLOTYPE: Diploidy
; TISSUE TYPE: Ovary
; CELL TYPE: Oocyte
; FEATURE:
; NAME/KEY: CDS

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1766
US-08-484-596A-15

Query Match 76.4%; Score 16.8; DB 2; Length 1840;
Best Local Similarity 90.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAA 21
|||||
Db 1061 CACCTGAACGCTAAGCGGGA 1042

RESULT 5
US-08-480-150A-15/c
; Sequence 15, Application US/08480150A
; Patent No. 5989550
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,150A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/149,223
; FILING DATE: 09-NOV-1993
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Felis domesticus
; DEVELOPMENTAL STAGE: Juvenile
; HAPLOTYPE: Diploidy
; TISSUE TYPE: Ovary
; CELL TYPE: Oocyte
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: 57..1766
US-08-480-150A-15

Query Match 76.4%; Score 16.8; DB 2; Length 1840;
Best Local Similarity 90.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAA 21
|||||
Db 1061 CACCTGAACGCTAAGCGGGA 1042

RESULT 6

US-08-458-731-15/c
; Sequence 15, Application US/08458731
; Patent No. 6001599
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,731
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Felis domesticus
; DEVELOPMENTAL STAGE: Juvenile
; HAPLOTYPE: Diploidy
; TISSUE TYPE: Ovary
; CELL TYPE: Oocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1766
US-08-458-731-15

Query Match 76.4%; Score 16.8; DB 3; Length 1840;

Best Local Similarity 90.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAA 21
|||||
Db 1061 CACCTGAACGCTAAGCGGGA 1042

RESULT 7

US-08-149-223A-15/c
; Sequence 15, Application US/08149223A
; Patent No. 6027727
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Jeffrey D.
; APPLICANT: Hsu, Kuang T.
; APPLICANT: Podolski, Joseph S.
; TITLE OF INVENTION: Materials and Methods for Immunocontraception
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/149,223A
; FILING DATE: 09-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/012,990
; FILING DATE: 29-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/973,341
; FILING DATE: 09-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 31745
; TELEPHONE: 312/474-6653
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1840 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORGANISM: Felis domesticus
; DEVELOPMENTAL STAGE: Juvenile
; HAPLOTYPE: Diploidy
; TISSUE TYPE: Ovary
; CELL TYPE: Oocyte
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1766
US-08-149-223A-15

Query Match 76.4%; Score 16.8; DB 3; Length 1840;
Best Local Similarity 90.0%; Pred. No. 57;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAA 21

Db 1061 CACCTGAACGCTAAGCGGGA 1042
|||||

RESULT 8

US-09-621-976-18882/c
; Sequence 18882, Application US/09621976

; Patent No. 6639063

; GENERAL INFORMATION:

; APPLICANT: Dumas Milne Edwards, J.B.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENST.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; SOFTWARE: Patent.pm

; SEQ ID NO 18882

; LENGTH: 461

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-621-976-18882

Query Match 73.6%; Score 16.2; DB 4; Length 461;

Best Local Similarity 85.7%; Pred. No. 94;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22

|||||

Db 108 CACCTGAAGATTAGCAGAAC 88

RESULT 9

US-09-270-767-13955/c

; Sequence 13955, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 13955

; LENGTH: 1070

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

US-09-270-767-13955

Query Match 73.6%; Score 16.2; DB 4; Length 1070;

Best Local Similarity 85.7%; Pred. No. 1.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22

|||||

Db 150 CACTTGAAGATTAGCGGGAAC 130

RESULT 10

US-09-583-110-515/c

; Sequence 515, Application US/09583110

; Patent No. 6699703

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.

; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

; FILE REFERENCE: PATH00-07A

; CURRENT APPLICATION NUMBER: US/09/583,110

; CURRENT FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: US 09/107,433

; PRIOR FILING DATE: 1998-06-30

; PRIOR APPLICATION NUMBER: US 60/085,131

; PRIOR FILING DATE: 1998-05-12

; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02

; NUMBER OF SEQ ID NOS: 5322

; SEQ ID NO 515

; LENGTH: 1104

; TYPE: DNA

; ORGANISM: Streptococcus pneumoniae

US-09-583-110-515

Query Match 73.6%; Score 16.2; DB 4; Length 1104;

Best Local Similarity 85.7%; Pred. No. 1.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22

|||||

Db 709 CACCCCAACGATAAGCTGAAC 689

RESULT 11

US-09-107-433-2176/c

; Sequence 2176, Application US/09107433

; Patent No. 6800744

; GENERAL INFORMATION:

; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID

; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE

; THERAPEUTICS

; NUMBER OF SEQUENCES: 5206

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: GENOME THERAPEUTICS CORPORATION

; STREET: 100 Beaver Street

; CITY: Waltham

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02354

; COMPUTER READABLE FORM:

; MEDIUM TYPE: CD-ROM ISO9660

; COMPUTER: <Unknown>

; OPERATING SYSTEM: <Unknown>

; SOFTWARE: <Unknown>

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/107,433

; FILING DATE: 30-Jun-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/ 085131

; FILING DATE: May 12, 1998

; APPLICATION NUMBER: 60/051553

; FILING DATE: July 2, 1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Ariniello, Pamela Deneke

; REGISTRATION NUMBER: 40,489

; REFERENCE/DOCKET NUMBER: GTC-011

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (781)893-5007

; TELEFAX: (781)893-8277

; INFORMATION FOR SEQ ID NO: 2176:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1128 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: circular

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Streptococcus pneumoniae

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (B) LOCATION 1...1128

; SEQUENCE DESCRIPTION: SEQ ID NO: 2176:

US-09-107-433-2176

```
Query Match 73.6%; Score 16.2; DB 4; Length 1128;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
||||| ||||||| |||||
DB 733 CACCCCAACGATAAGCTGAAC 713

RESULT 12
US-08-215-709-2/c
; Sequence 2, Application US/08215709
; Patent No. 5432071
; GENERAL INFORMATION:
; APPLICANT: ICHIKAWA, Toshio
; APPLICANT: KOYAMA, Yasuji
; APPLICANT: OTAKE, Hideo
; APPLICANT: NAKANO, Eiichi
; TITLE OF INVENTION: Variant E1 Protein Gene For Pyruvate
; TITLE OF INVENTION: Dehydrogenase Complex And Variant E1 Protein Of Pyruvate
; TITLE OF INVENTION: Dehydrogenase Complex
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 22-MAR-1994
; APPLICATION NUMBER: US/08/215,709
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jean B. Fordis
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 04853.0011-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2664 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-215-709-2

Query Match 73.6%; Score 16.2; DB 1; Length 2664;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
||||| ||||||| |||||
DB 2376 CACCTGAGCGATATACGGAAC 2356

RESULT 13
US-08-961-527-3/c
; Sequence 3, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340PI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-961-527-3

Query Match 73.6%; Score 16.2; DB 3; Length 26385;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
||||| ||||||| |||||
DB 18002 CACCCCAACGATAAGCTGAAC 17982

RESULT 14
US-09-949-016-16982
; Sequence 16982, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16982
; LENGTH: 51967
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(51967)
; OTHER INFORMATION: n = A,T,C or G
; US-09-949-016-16982

Query Match 73.6%; Score 16.2; DB 4; Length 51967;
```

Best Local Similarity 85.7%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCGGAA 21
|||
DB 27567 CCTCCTGAATGATAGAGGAA 27587

RESULT 15
US-10-061-943A-1/c
; Sequence 1, Application US/10061943A
; Patent No. 6740324
; GENERAL INFORMATION:
; APPLICANT: Schall, Thomas J.
; APPLICANT: Penfold, Mark E.T.
; APPLICANT: ChemoCentryx, Inc.
; TITLE OF INVENTION: Methods and Compositions Useful for Stimulating an
; TITLE OF INVENTION: Immune Response
; FILE REFERENCE: 019934-001610US
; CURRENT APPLICATION NUMBER: US/10/061.943A
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/265,925
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Rhesus cytomegalovirus
; FEATURE:
; OTHER INFORMATION: rhesus monkey (Macaca mulatta) cytomegalovirus
; OTHER INFORMATION: (rhCMV) short unique region 28.1 (rhUS28.1) coding
; OTHER INFORMATION: sequence
US-10-061-943A-1

Query Match 71.8%; Score 15.8; DB 4; Length 1020;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATAAGCGGAAC 22
|||||
DB 313 CCTGAACGATAAGCGGAAC 295

RESULT 16
US-09-799-451-459
; Sequence 459, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yunqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948

RESULT 18
US-09-188-082-8/c
; Sequence 8, Application US/09188082
; Patent No. 6270765
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/188,082
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-188-082-8
Query Match 70.9%; Score 15.6; DB 3; Length 53;
Best Local Similarity 81.8%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CCACCTGAACGATAGCGGAAC 22
Db 32 CTACTTGAACCTTAGCGGAAC 11
RESULT 19
US-09-364-088-8/c
; Sequence 8, Application US/09364088
; Patent No. 6365161
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo, et al.
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street, 24th Floor
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/364,088
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/188,082
; FILING DATE: 07-JUNE-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,172
; FILING DATE: 07-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remillard, Jane E.
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: MXI-043CP2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-7414
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-364-088-8
Query Match 70.9%; Score 15.6; DB 3; Length 53;
Best Local Similarity 81.8%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 CCACCTGAACGATAGCGGAAC 22
Db 32 CTACTTGAACCTTAGCGGAAC 11
RESULT 20
US-09-102-716-8/c
; Sequence 8, Application US/09102716
; Patent No. 6395272
; GENERAL INFORMATION:
; APPLICANT: Yashwant M. Deo
; APPLICANT: Joel Goldstein
; APPLICANT: Robert Graziano
; APPLICANT: Chezian Somasundaram
; TITLE OF INVENTION: THERAPEUTIC COMPOUNDS COMPRISED
; OF ANTI-FC RECEPTOR ANTIBODIES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/102,716
; FILING DATE: 22-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/661,052
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MXI-043CP
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 53 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-102-716-8

Query Match 70.9%; Score 15.6; DB 3; Length 53;
Best Local Similarity 81.8%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
|||
DB 32 CTACTTGAACCTTAAAGCGGAAC 11
|||

RESULT 21

US-09-489-039A-1339/c
; Sequence 1339, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1339
; LENGTH: 585
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1339

Query Match 70.9%; Score 15.6; DB 4; Length 585;
Best Local Similarity 81.8%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
|||
DB 529 CGATCTGAACGACACGCGGAAC 508
|||

RESULT 22

US-09-489-039A-95/c
; Sequence 95, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 95
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-95

Query Match 70.9%; Score 15.6; DB 4; Length 594;
Best Local Similarity 81.8%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
|||
DB 527 CCACCTGCAGGATACGCCGAAC 506
|||

RESULT 23

US-08-998-416-821/c
; Sequence 821, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:

APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCES/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 821:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1522RP
US-08-998-416-821

Query Match 70.9%; Score 15.6; DB 3; Length 705;
Best Local Similarity 81.8%; Pred. No. 2e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
|||
DB 638 CCACGTGACCGACAAGGGGAAC 617
|||

RESULT 24

US-09-489-039A-101
; Sequence 101, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

APPLICANT: Gary Breton et. al


```
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 1432
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-1432

Query Match          70.9%; Score 15.6; DB 4; Length 1398;
Best Local Similarity 81.8%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 CCACCTGAACGATAAGCGGAAC 22
Db      1203 CCATCTGAACGATACGCGGAGC 1224

RESULT 30
US-09-252-991A-6692
; Sequence 6692, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6692
; LENGTH: 1482
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6692

Query Match          70.9%; Score 15.6; DB 4; Length 1482;
Best Local Similarity 81.8%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 CCACCTGAACGATAAGCGGAAC 22
Db      846 CCACCTGTACGATCCGCGAGAAC 867

RESULT 31
US-09-252-991A-6662
; Sequence 6662, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
```

```
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6662
; LENGTH: 1791
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6662

Query Match          70.9%; Score 15.6; DB 4; Length 1791;
Best Local Similarity 81.8%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 CCACCTGAACGATAAGCGGAAC 22
Db      151 CCACCTGTACGATCCGCGAGAAC 172

RESULT 32
US-09-902-540-3949
; Sequence 3949, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3949
; LENGTH: 2127
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-3949

Query Match          70.9%; Score 15.6; DB 4; Length 2127;
Best Local Similarity 81.8%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 CCACCTGAACGATAAGCGGAAC 22
Db      438 CCACCTGAACGATAAGCGGAAC 459

RESULT 33
US-09-902-540-1194/c
; Sequence 1194, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1194
; LENGTH: 21295
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1194

Query Match          70.9%; Score 15.6; DB 4; Length 21295;
Best Local Similarity 81.8%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

QY 1 CCACCTGAACGATAACGGGAAC 22
||| ||||| || ||||| |||||
Db 15102 CCACGCTGAAGGAGAACTGAAC 15081

RESULT 34
US-09-949-016-12735
; Sequence 12735, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12735
; LENGTH: 80355
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12735

Query Match 70.9%; Score 15.6; DB 4; Length 80355;
Best Local Similarity 81.8%; Pred. No. 4.3e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAACGGGAAC 22
||| ||||| || ||||| |||||
Db 22907 CCACCTGACCAATAAGCCAAAC 22928

RESULT 35
US-09-949-016-13572
; Sequence 13572, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13572
; LENGTH: 80357
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13572

Query Match 70.9%; Score 15.6; DB 4; Length 80357;
Best Local Similarity 81.8%; Pred. No. 4.3e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAACGGGAAC 22
||| ||||| || ||||| |||||
Db 22907 CCACCTGACCAATAAGCCAAAC 22928

RESULT 36
US-09-596-002-35
; Sequence 35, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Patterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; CURRENT FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 96109
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 35
; PUBLICATION INFORMATION:
US-09-596-002-35

Query Match 70.9%; Score 15.6; DB 4; Length 96109;
Best Local Similarity 81.8%; Pred. No. 4.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAACGGGAAC 22
||| ||||| || ||||| |||||
Db 94885 CCACCTGAACTAAAGCAGATC 94906

RESULT 37
US-09-540-236-1876/c
; Sequence 1876, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1876
; LENGTH: 765
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-1876

Query Match 70.0%; Score 15.4; DB 4; Length 765;
Best Local Similarity 94.1%; Pred. No. 2.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCACCTGAACGATAACGG 18
||| ||||| ||||| |||||
Db 184 CACCAGAACGATAACGG 168

RESULT 38
US-09-602-777A-359/c
; Sequence 359, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor

;; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;; FILE REFERENCE: INVOLVED IN HOMEOSTASIS AND ADAPTATION
;; CURRENT FILING DATE: 2000-06-23

;; PRIOR APPLICATION NUMBER: US 60/141031
;; PRIOR FILING DATE: 1999-06-25

;; PRIOR APPLICATION NUMBER: DE 19931636.8
;; PRIOR FILING DATE: 1999-07-08

;; PRIOR APPLICATION NUMBER: DE 19932125.6
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932126.4
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932127.2
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932128.0
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932129.9
;; PRIOR FILING DATE: 1999-07-19

;; PRIOR APPLICATION NUMBER: DE 19932226.0
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932920.6
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932922.2
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932924.9
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932928.1
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932930.3
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932933.8
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932935.4
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932973.7
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19933002.6
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19933003.4
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19933005.0
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19933006.9
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19941378.9
;; PRIOR FILING DATE: 1999-08-31

;; PRIOR APPLICATION NUMBER: DE 19941379.7
;; PRIOR FILING DATE: 1999-08-31

;; PRIOR APPLICATION NUMBER: DE 19941390.8
;; PRIOR FILING DATE: 1999-08-31

;; PRIOR APPLICATION NUMBER: DE 19941391.6
;; PRIOR FILING DATE: 1999-08-31

;; PRIOR APPLICATION NUMBER: DE 19942088.2
;; PRIOR FILING DATE: 1999-09-03

;; NUMBER OF SEQ ID NOS: 442
;; SEQ ID NO 359
;; LENGTH: 3945

;; TYPE: DNA
;; ORGANISM: Corynebacterium glutamicum

;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (101)..(3922)

;; OTHER INFORMATION: RXN01499
US-09-602-777A-359

Query Match 70.0%; Score 15.4; DB 4; Length 3945;
Best Local Similarity 94.1%; Pred. No. 3.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CACCTGAACGATAAGCG 18
Db 1169 CACCGAACGATAAGCG 1153

;; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
;; FILE REFERENCE: INVOLVED IN HOMEOSTASIS AND ADAPTATION
;; CURRENT FILING DATE: 2000-06-23

;; PRIOR APPLICATION NUMBER: US 60/141031
;; PRIOR FILING DATE: 1999-06-25

;; PRIOR APPLICATION NUMBER: DE 19931636.8
;; PRIOR FILING DATE: 1999-07-08

;; PRIOR APPLICATION NUMBER: DE 19932125.6
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932126.4
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932127.2
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932128.0
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932129.9
;; PRIOR FILING DATE: 1999-07-19

;; PRIOR APPLICATION NUMBER: DE 19932226.0
;; PRIOR FILING DATE: 1999-07-09

;; PRIOR APPLICATION NUMBER: DE 19932920.6
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932922.2
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932924.9
;; PRIOR FILING DATE: 1999-07-14

;; PRIOR APPLICATION NUMBER: DE 19932928.1
;; PRIOR FILING DATE: 1999-07-14

RESULT 39

US-09-596-002-41/c

;; Sequence 41, Application US/09596002

;; Patent No. 6632636

;; GENERAL INFORMATION:

;; APPLICANT: Lagace, Robert, E.

;; APPLICANT: Patterson, Chandra

;; APPLICANT: Berg, Kim, L.

;; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME

;; FILE REFERENCE: PM-0008-4 US

;; CURRENT APPLICATION NUMBER: US/09/596,002

;; CURRENT FILING DATE: 2000-06-16

;; PRIOR APPLICATION NUMBER: 60/140,121

;; PRIOR FILING DATE: 1999-06-18

;; NUMBER OF SEQ ID NOS: 41

;; SOFTWARE: PERL Program

;; SEQ ID NO 41

;; LENGTH: 269223

;; TYPE: DNA

;; ORGANISM: Moraxella catarrhalis

;; FEATURE:

;; NAME/KEY: misc feature

;; OTHER INFORMATION: Incyte template ID No. 6632636 41

;; PUBLICATION INFORMATION:

US-09-596-002-41

Query Match 70.0%; Score 15.4; DB 4; Length 269223;

Best Local Similarity 94.1%; Pred. No. 6.3e+02;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CACCTGAACGATAAGCG 18

Db 18929 CACCGAACGATAAGCG 18913

;; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

;; FILE REFERENCE: File Reference: 7326-094

;; CURRENT APPLICATION NUMBER: US/09/270,767

;; CURRENT FILING DATE: 1999-03-17

;; NUMBER OF SEQ ID NOS: 62517

;; SOFTWARE: Patentin Ver. 2.0

;; SEQ ID NO 1176

;; LENGTH: 662

;; TYPE: DNA

;; ORGANISM: Drosophila melanogaster

US-09-270-767-1176

Query Match 69.1%; Score 15.2; DB 4; Length 662;

Best Local Similarity 85.0%; Pred. No. 3.2e+02;

Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CACCTGAACGATAAGCGAA 21

Db 614 CCCCTGAAAGTTAAGCGAA 595

Search completed: June 4, 2005, 11:52:56

Job time : 50.9467 secs

This Page Blank (uspto)


```
US-10-270-333-148
; Sequence 148, Application US/10270333
; Publication No. US20030092124A1
; GENERAL INFORMATION:
; APPLICANT: Cravchik, Anibal
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF AS INSECTICIDAL TARGETS
; FILE REFERENCE: CL000733CON
; CURRENT APPLICATION NUMBER: US/10/270,333
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148
; LENGTH: 3772
; TYPE: DNA
; ORGANISM: Drosophila
US-10-270-333-148

Query Match          78.2%; Score 17.2; DB 14; Length 3772;
Best Local Similarity 86.4%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCACCTGAACGATAAGCGGAAC 22
Db 845 CAATCTGAACGATAAGCGGCAC 866

RESULT 3
US-10-741-849-6229
; Sequence 6229, Application US/10741849
; Publication No. US20050019931A1
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Nucleic Acids Encoding Anti-fungal Drug Targets and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: 10182-023-999
; CURRENT APPLICATION NUMBER: US/10/741,849
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/434,832
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 8000
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6229
; LENGTH: 6051
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-741-849-6229

Query Match          78.2%; Score 17.2; DB 19; Length 6051;
Best Local Similarity 86.4%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCACCTGAACGATAAGCGGAAC 22
Db 2857 CCACCTCAACACACAGCGGGAAC 2878

RESULT 4
US-10-451-467A-667
; Sequence 667, Application US/10451467A
; Publication No. US20040161840A1
; GENERAL INFORMATION:
; APPLICANT: CONTRERAS, ROLAND HENRI
```

```
; APPLICANT: EBERHARDT, INES
; APPLICANT: LUYTEN, WALTER HERMAN MARIA LOUIS
; APPLICANT: REEKMAN, RIEKA JOSEPHINA
; TITLE OF INVENTION: BAX-RESPONSIVE GENES FOR DRUG TARGET IDENTIFICATION IN
; TITLE OF INVENTION: YEAST AND FUNGI
; FILE REFERENCE: JAB-1667
; CURRENT APPLICATION NUMBER: US/10/451,467A
; CURRENT FILING DATE: 2003-06-19
; PRIOR APPLICATION NUMBER: EP 00870318.3
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: EP 01870002.1
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: EP 01870003.9
; PRIOR FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 732
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 667
; LENGTH: 8496
; TYPE: DNA
; ORGANISM: Candida albicans
US-10-451-467A-667

Query Match          78.2%; Score 17.2; DB 18; Length 8496;
Best Local Similarity 86.4%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCACCTGAACGATAAGCGGAAC 22
Db 6726 CCACCTCAACACACAGCGGGAAC 6747

RESULT 5
US-10-437-963-67282
; Sequence 67282, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 67282
; LENGTH: 3410
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_68153C.1
US-10-437-963-67282

Query Match          76.4%; Score 16.8; DB 18; Length 3410;
Best Local Similarity 90.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCACCTGAACGATAAGCGGA 20
Db 389 CCATCTGAACCATAGCGGA 408

RESULT 6
US-10-472-724-5
; Sequence 5, Application US/10472724
; Publication No. US20040171806A1
; GENERAL INFORMATION:
; APPLICANT: Cid-Arregui, Angel
```

```
; APPLICANT: Zur Hausen, Harald
; TITLE OF INVENTION: Modified HPV E6 and E7 genes and proteins useful for vaccination
; FILE REFERENCE: 4121-154
; CURRENT APPLICATION NUMBER: US/10/472,724
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/EP02/03271
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: EP 01107271.7
; PRIOR FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutagenized E6 gene of HPV-18
; NAME/KEY: CDS
; LOCATION: (14)..(529)
US-10-472-724-5

Query Match          74.5%; Score 16.4; DB 18; Length 545;
Best Local Similarity 94.4%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCG 18
    |||||
Db 388 CCACCTGAACGAGAGCG 405

RESULT 7
US-10-424-599-77243
; Sequence 77243, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 77243
; LENGTH: 373
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_40766C.1
US-10-424-599-77243

Query Match          73.6%; Score 16.2; DB 17; Length 373;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGAAC 22
    |||||
Db 89 CACCTGAACGAGATGAGAAC 109

RESULT 8
US-10-487-901-6984/c
; Sequence 6984, Application US/10487901
; Publication No. US20050091708A1
; GENERAL INFORMATION:
; APPLICANT: Oreido, Jeremiah Vincent
; APPLICANT: McCreery, David
; APPLICANT: Pell, Randy
; APPLICANT: Miller, Barbara
; APPLICANT: Weglarz, Thaddeus
```

```
; APPLICANT: Gachotte, Daniel
; APPLICANT: Blakeslee, Beth
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Reddy, Avutu
; APPLICANT: Shukla, Vipula
; APPLICANT: Croasley, Rodney
; TITLE OF INVENTION: Nucleic Acid Compositions Conferring Altered Metabolic Character
; FILE REFERENCE: DOW-08552
; CURRENT APPLICATION NUMBER: US/10/487,901
; CURRENT FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7560
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6984
; LENGTH: 730
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-487-901-6984

Query Match          73.6%; Score 16.2; DB 19; Length 730;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGAAC 22
    |||||
Db 292 CACCTGAACCGTAAGCGAAC 272

RESULT 9
US-10-437-963-87359
; Sequence 87359, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87359
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86311C.1
US-10-437-963-87359

Query Match          73.6%; Score 16.2; DB 18; Length 804;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCGAA 21
    |||||
Db 579 CGACCTGAACGCTAAGCTGA 599

RESULT 10
US-10-425-115-132310
; Sequence 132310, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

```
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 132310
; LENGTH: 829
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_52152C.1
US-10-425-115-132310

Query Match          73.6%; Score 16.2; DB 18; Length 829;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAAGCGGAAC 22
        ||||| ||||| ||||| |||||
Db      510 CACCGGATCTATAAGCGGAAC 530

RESULT 11
US-10-437-963-40152
; Sequence 40152, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 40152
; LENGTH: 927
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43622C.1
US-10-437-963-40152

Query Match          73.6%; Score 16.2; DB 18; Length 927;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAAGCGGAAC 22
        ||||| ||||| ||||| |||||
Db      847 CACCTCAAGGATAAGAGGAAC 867

RESULT 12
US-09-938-842A-819/c
; Sequence 819, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
```

```
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 819
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-819

Query Match          73.6%; Score 16.2; DB 9; Length 1005;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAAGCGGAAC 22
        ||||| ||||| ||||| |||||
Db      430 CACCTGAACCGTAAGCGGAAC 410

RESULT 13
US-09-938-842A-819/c
; Sequence 819, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 819
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-819

Query Match          73.6%; Score 16.2; DB 11; Length 1005;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAAGCGGAAC 22
        ||||| ||||| ||||| |||||
Db      430 CACCTGAACCGTAAGCGGAAC 410

RESULT 14
US-10-472-928-3471/c
; Sequence 3471, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON Spa
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926W0
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
```

```
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3471
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-472-928-3471

Query Match          73.6%; Score 16.2; DB 19; Length 1101;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAGCGGAAC 22
      ||||| ||||| ||||| |||||
Db      709  CACCCCAACGATAGCTGAAC 689

RESULT 15
US-10-282-122A-37938/c
; Sequence 37938, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37938
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-282-122A-37938

Query Match          73.6%; Score 16.2; DB 17; Length 1104;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAGCGGAAC 22
      ||||| ||||| ||||| |||||
Db      709  CACCCCAACGATAGCTGAAC 689

; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3471
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-472-928-3471

Query Match          73.6%; Score 16.2; DB 19; Length 1101;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAGCGGAAC 22
      ||||| ||||| ||||| |||||
Db      709  CACCCCAACGATAGCTGAAC 689

RESULT 16
US-10-425-114-14697/c
; Sequence 14697, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 14697
; LENGTH: 1349
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB24-134-C12_FLI
US-10-425-114-14697

Query Match          73.6%; Score 16.2; DB 17; Length 1349;
Best Local Similarity 85.7%; Pred. No. 5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      2  CACCTGAACGATAGCGGAAC 22
      ||||| ||||| ||||| |||||
Db      493  CACCTGAACGTAAGCGGAAC 473

RESULT 17
US-09-934-455-291/c
; Sequence 291, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pineda, Omaira
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 291
; LENGTH: 1699
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (417)..(1421)
; OTHER INFORMATION: G5
```

US-09-934-455-291

Query Match 73.6%; Score 16.2; DB 10; Length 1699;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
||||||| ||||| ||||| ||||| |||||
Db 846 CACCTGAACCGTAAGCGGAAC 826

RESULT 18

US-10-278-173-81/c
; Sequence 81, Application US/10278173
; Publication No. US20030061637A1

; GENERAL INFORMATION:

; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Broun, Pierre
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Pineda, Omaira
; APPLICANT: Zhang, James
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Pilgrim, Marsha
; APPLICANT: Keddie, James
; APPLICANT: Heard, Jacqueline
; APPLICANT: Reuber, Lynne
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc
; APPLICANT: Samaha, Raymond

; TITLE OF INVENTION: POLYNUCLEOTIDES FOR ROOT TRAIT ALTERATION

; FILE REFERENCE: MBI-009

; CURRENT APPLICATION NUMBER: US/10/278,173

; CURRENT FILING DATE: 2002-10-21

; PRIOR APPLICATION NUMBER: US/09/533,392

; PRIOR FILING DATE: 2000-03-22

; PRIOR APPLICATION NUMBER: 60/125,814

; PRIOR FILING DATE: 1999-03-23

; NUMBER OF SEQ ID NOS: 177

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 81

; LENGTH: 1699

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; OTHER INFORMATION: G5

US-10-278-173-81

Query Match 73.6%; Score 16.2; DB 14; Length 1699;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
||||||| ||||| ||||| ||||| |||||
Db 846 CACCTGAACCGTAAGCGGAAC 826

RESULT 19

US-10-295-403-3/c

; Sequence 3, Application US/10295403

; Publication No. US2003010481A1

; GENERAL INFORMATION:

; APPLICANT: Heard, Jacqueline
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc
; APPLICANT: Broun, Pierre
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddie, James
; APPLICANT: Zhang, James
; APPLICANT: Benito, Maria-Ines
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Fromm, Mike

; TITLE OF INVENTION: PLANT GENE SEQUENCES I

; FILE REFERENCE: MBI-0003

; CURRENT APPLICATION NUMBER: US/10/295,403

; CURRENT FILING DATE: 2002-11-15

; PRIOR APPLICATION NUMBER: US/09/394,519

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: 60/101,349

; PRIOR FILING DATE: 1998-09-22

; PRIOR APPLICATION NUMBER: 60/103,312

; PRIOR FILING DATE: 1998-10-06

; PRIOR APPLICATION NUMBER: 60/108,734

; PRIOR FILING DATE: 1998-11-17

; PRIOR APPLICATION NUMBER: 60/113,409

; PRIOR FILING DATE: 1998-12-22

; NUMBER OF SEQ ID NOS: 170

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1699

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (417)..(1418)

; OTHER INFORMATION: G5

US-10-295-403-3

Query Match 73.6%; Score 16.2; DB 15; Length 1699;

Best Local Similarity 85.7%; Pred. No. 5.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
||||||| ||||| ||||| ||||| |||||
Db 846 CACCTGAACCGTAAGCGGAAC 826

RESULT 20

US-10-225-066A-721/c

; Sequence 721, Application US/10225066A

; Publication No. US20030226173A1

; GENERAL INFORMATION:

; APPLICANT: Mendel Biotechnology, Inc.

; APPLICANT: Ratcliffe, Oliver

; APPLICANT: Riechmann, Jose Luis

; APPLICANT: Adam, Luc J

; APPLICANT: DUBELL, Arnold T

; APPLICANT: HEARD, Jacqueline E

; APPLICANT: PILGRIM, Marsha L

; APPLICANT: JIANG, Cai-Zhong

; APPLICANT: REUBER, T. Lynne

; APPLICANT: CREELMAN, Robert A

; APPLICANT: PINEDA, Omaira

; APPLICANT: YU, Guo-Liang

; APPLICANT: BROWN, Pierre E

; TITLE OF INVENTION: Yield-Related Polynucleotides and Polypeptides in Plants

; FILE REFERENCE: MBI0036-2 US

; CURRENT APPLICATION NUMBER: US/10/225,066A

; CURRENT FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: 09/837,444

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/310,847

; PRIOR FILING DATE: 2001-08-09

; PRIOR APPLICATION NUMBER: 60/336,049

; PRIOR FILING DATE: 2001-12-05

; PRIOR APPLICATION NUMBER: 60/338,692

; PRIOR FILING DATE: 2001-12-11

; PRIOR APPLICATION NUMBER: 10/171,468

; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 1122

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 721

; LENGTH: 1699

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana


```

; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1711
; LENGTH: 1699
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-412-699B-1711
```

```

Query Match 73.6%; Score 16.2; DB 17; Length 1699;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```

QY 2 CACCTGAACGATAGCGGAC 22
||||||| ||||| |||
Db 846 CACCTGAACCGTAAGCGAAC 826
```

```

RESULT 24
US-10-425-114-13035/c
; Sequence 13035, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
```

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
```

```

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 13035
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana columbia
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-ATXLIB327432P4F01_FLI
US-10-425-114-13035
```

```

Query Match 73.6%; Score 16.2; DB 17; Length 1813;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```

QY 2 CACCTGAACGATAGCGGAC 22
||||||| ||||| |||
Db 961 CACCTGAACCGTAAGCGAAC 941
```

RESULT 25

```

US-10-282-122A-12230
; Sequence 12230, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12230
; LENGTH: 1863
; TYPE: DNA
; ORGANISM: Bacteroides fragilis
US-10-282-122A-12230
```

```

Query Match 73.6%; Score 16.2; DB 17; Length 1863;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
```

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCGGAA 21
 |||||
 Db 1635 CCGCCTGAACGATAAGCGGAA 1655

RESULT 26
 US-10-369-493-33470/c
 ; Sequence 33470, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 33470
 ; LENGTH: 2164
 ; TYPE: DNA
 ; ORGANISM: Desulfitobacterium hafnienae
 US-10-369-493-33470

Query Match 73.6%; Score 16.2; DB 17; Length 2164;
 Best Local Similarity 85.7%; Pred. No. 5.2e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCGGAA 21
 |||||
 Db 1393 CCACATGAACGATAAGCGGAA 1373

RESULT 27
 US-10-437-963-63022
 ; Sequence 63022, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 63022
 ; LENGTH: 2694
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_64304C.1
 US-10-437-963-63022

Query Match 73.6%; Score 16.2; DB 18; Length 2694;
 Best Local Similarity 85.7%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCGGAA 21
 |||||
 Db 1393 CCACATGAACGATAAGCGGAA 1373

Db 729 CGACCTGAAGAGAGCGGAA 749

RESULT 28
 US-10-437-963-63015
 ; Sequence 63015, Application US/10437963
 ; Publication No. US20040123343A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Wu, Wei
 ; APPLICANT: Boukharov, Andrey A.
 ; APPLICANT: Barbazuk, Brad
 ; APPLICANT: Li, Ping
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: 38-21(53221)B
 ; CURRENT APPLICATION NUMBER: US/10/437,963
 ; CURRENT FILING DATE: 2003-05-14
 ; NUMBER OF SEQ ID NOS: 204966
 ; SEQ ID NO 63015
 ; LENGTH: 2946
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT4530_64299C.1
 US-10-437-963-63015

Query Match 73.6%; Score 16.2; DB 18; Length 2946;
 Best Local Similarity 85.7%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCGGAA 21
 |||||
 Db 729 CGACCTGAAGAGAGCGGAA 749

RESULT 29
 US-10-893-671-3/c
 ; Sequence 3, Application US/10893671
 ; Publication No. US20050064527A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Levy, Stuart, et. al.
 ; TITLE OF INVENTION: NIMR COMPOSITIONS AND THEIR METHODS OF USE
 ; FILE REFERENCE: PKZ-043
 ; CURRENT APPLICATION NUMBER: US/10/893,671
 ; CURRENT FILING DATE: 2004-07-15
 ; PRIOR APPLICATION NUMBER: US/09/801,563
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 60/188,362
 ; PRIOR FILING DATE: 2000-03-10
 ; NUMBER OF SEQ ID NOS: 98
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 10157
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (5397)..(8060)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (8075)..(9967)
 US-10-893-671-3

Query Match 73.6%; Score 16.2; DB 19; Length 10157;
 Best Local Similarity 85.7%; Pred. No. 6e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCACCTGAACGATAAGCGGAA 22
 |||||
 Db 729 CGACCTGAAGAGAGCGGAA 749

Db 7772 CACCTGAGCATATACGGAAC 7752

RESULT 30
US-08-961-527-3/c
; Sequence 3, Application US/08961527
; Publication No. US20020032323A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear

US-08-961-527-3

Query Match 73.6%; Score 16.2; DB 8; Length 26385;
Best Local Similarity 85.7%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGACGATAAGCGGAAC 22
|||||
Db 18002 CACCCCAACGATAAGCTGAAC 17982

RESULT 31
US-158-844-3/c
; Sequence 3, Application US/10158844
; Publication No. US20040029118A1
; GENERAL INFORMATION:
; APPLICANT: Kunsch et al.
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude Pentium 3
; OPERATING SYSTEM: Windows 98

; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/158,844
; FILING DATE: 03-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/961,527
; FILING DATE: 1997-10-30
; APPLICATION NUMBER: US 60/029,960
; FILING DATE: 1996-10-31
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB340P1D1
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26385 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-158-844-3

Query Match 73.6%; Score 16.2; DB 17; Length 26385;
Best Local Similarity 85.7%; Pred. No. 6.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
|||||
Db 18002 CACCCCAACGATAAGCTGAAC 17982

RESULT 32
US-10-472-928-4979
; Sequence 4979, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; PRIOR FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 4979
; LENGTH: 2162598
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-10-472-928-4979

Query Match 73.6%; Score 16.2; DB 19; Length 2162598;
Best Local Similarity 85.7%; Pred. No. 8.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
|||||
Db 1480095 CACCCCAACGATAAGCTGAAC 1480115

RESULT 33
US-09-918-995-32213
; Sequence 32213, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32213
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(493)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-32213

Query Match 71.8%; Score 15.8; DB 10; Length 493;
Best Local Similarity 89.5%; Pred. No. 7.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATAAGCGGAAC 22
| | | | | | | | | | | | | | | | | | | | | |
Db 235 CATGAACGAGAGCGGAAC 253

RESULT 34

US-10-424-599-71212/c
; Sequence 71212, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 71212
; LENGTH: 507
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(507)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_35316C.1
US-10-424-599-71212

Query Match 71.8%; Score 15.8; DB 17; Length 507;
Best Local Similarity 89.5%; Pred. No. 7.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAAGCGGAA 21
| | | | | | | | | | | | | | | | | | | | | |
Db 40 ACCTGAACGAGAGCTGAA 22

RESULT 35

US-10-369-493-33989/c
; Sequence 33989, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33989
; LENGTH: 873
; TYPE: DNA
; ORGANISM: Cytophaga hutchinsonii
US-10-369-493-33989

Query Match 71.8%; Score 15.8; DB 17; Length 873;
Best Local Similarity 89.5%; Pred. No. 7.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATAAGCGGAAC 22
| | | | | | | | | | | | | | | | | | | | | |
Db 769 CCTGAACGATCAGCGGAAC 751

RESULT 36

US-10-369-493-35273/c
; Sequence 35273, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 35273
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-35273

Query Match 71.8%; Score 15.8; DB 17; Length 1011;
Best Local Similarity 89.5%; Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATAAGCGGAAC 22
| | | | | | | | | | | | | | | | | | | | | |
Db 328 CCGAACGATCAGCGGAAC 310

RESULT 37

US-10-369-493-38543/c
; Sequence 38543, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38543
; LENGTH: 1011
; TYPE: DNA

; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38543

Query Match 71.8%; Score 15.8; DB 17; Length 1011;
Best Local Similarity 89.5%; Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATACGGGAAC 22
||| ||||| ||||| |||||
DB 328 CCGGAACGATCAGCGGAAC 310

RESULT 38

US-10-369-493-38732/c
; Sequence 38732, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 38732
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Agrobacterium tumefaciens
US-10-369-493-38732

Query Match 71.8%; Score 15.8; DB 17; Length 1011;
Best Local Similarity 89.5%; Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATACGGGAAC 22
||| ||||| ||||| |||||
DB 328 CCGGAACGATCAGCGGAAC 310

RESULT 39

US-09-944-049-5/c
; Sequence 5, Application US/09944049
; Publication No. US20030175681A1
; GENERAL INFORMATION:
; APPLICANT: Schall, Thomas J.
; APPLICANT: Penfold, Mark E.T.
; APPLICANT: ChemoCentryx, Inc.
; TITLE OF INVENTION: Inhibition of CMV Infection and Dissemination
; FILE REFERENCE: 019934-002510US
; CURRENT APPLICATION NUMBER: US/09/944,049
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/229,365
; PRIOR FILING DATE: 2000-08-30
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Rhesus cytomegalovirus
; FEATURE:
; OTHER INFORMATION: rhesus monkey cytomegalovirus (rhCMV) US28 homolog
; NAME/KEY: CDS
; LOCATION: (1)..(1020)
; OTHER INFORMATION: rhUS28.1
US-09-944-049-5

Query Match 71.8%; Score 15.8; DB 10; Length 1020;
Best Local Similarity 89.5%; Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATACGGGAAC 22
||| ||||| ||||| |||||
DB 313 CCTGAACGAAAGGGGAAC 295

RESULT 40

US-10-061-943A-1/c
; Sequence 1, Application US/10061943A
; Publication No. US20020176870A1
; GENERAL INFORMATION:
; APPLICANT: Schall, Thomas J.
; APPLICANT: Penfold, Mark E.T.
; APPLICANT: ChemoCentryx, Inc.
; TITLE OF INVENTION: Methods and Compositions Useful for Stimulating an
; TITLE OF INVENTION: Immune Response
; FILE REFERENCE: 019934-001610US
; CURRENT APPLICATION NUMBER: US/10/061,943A
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/265,925
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Rhesus cytomegalovirus
; FEATURE:
; OTHER INFORMATION: rhesus monkey (Macaca mulatta) cytomegalovirus
; OTHER INFORMATION: (rhCMV) short unique region 28.1 (rhUS28.1) coding
; OTHER INFORMATION: sequence
US-10-061-943A-1

Query Match 71.8%; Score 15.8; DB 13; Length 1020;
Best Local Similarity 89.5%; Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 CCTGAACGATACGGGAAC 22
||| ||||| ||||| |||||
DB 313 CCTGAACGAAAGGGGAAC 295

Search completed: June 4, 2005, 12:19:56
Job time : 213.716 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 4, 2005, 06:32:00 ; Search time 1343.17 Seconds
(without alignments)
623.460 Million cell updates/sec

Title: US-09-674-277-21
Perfect score: 22
Sequence: 1 ccacctgaacgataacggaac 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17.8	80.9	297	7 W66011	W66011 TGESTzy7le0
C 2	17.8	80.9	396	4 BJ434099	BJ434099 BJ434099
C 3	17.8	80.9	547	4 BM176732	BM176732 TGESTzya9
C 4	17.8	80.9	701	6 CB028481	CB028481 TGESTzyd8
C 5	17.4	79.1	102	8 CC180655	CC180655 O180568-0
C 6	17.4	79.1	529	4 BM774282	BM774282 r391e12.y
C 7	17.4	79.1	1969	9 CG754242	CG754242 P049-3-E0
C 8	17.2	78.2	329	2 BB378617	BB378617 BB378617
C 9	17.2	78.2	417	8 AQ173857	AQ173857 HS 3199 B
C 10	17.2	78.2	582	6 CA067617	CA067617 SCQSD105
C 11	17.2	78.2	587	1 AU022955	AU022955 AU022955
C 12	17.2	78.2	739	9 CL515918	CL515918 SAIL_908
C 13	17.2	78.2	1081	5 BQ053756	BQ053756 AGENCOURT
C 14	17	77.3	615	8 AQ325117	AQ325117 mgxb0020H
C 15	17	77.3	1038	9 CNS014C4	AL103966 Drosophila
C 16	16.8	76.4	407	2 BF742895	BF742895 IL2-BF073
C 17	16.8	76.4	425	8 BZ291938	BZ291938 SALK_1227
C 18	16.8	76.4	578	9 CR341213	CR341213 Medicago
C 19	16.8	76.4	634	1 AI398989	AI398989 GH21181.5
C 20	16.8	76.4	637	4 BI363699	BI363699 RE48622.5
C 21	16.8	76.4	649	1 AA979615	AA979615 LD34308.5
C 22	16.8	76.4	651	4 BI229963	BI229963 RE27868.5
C 23	16.8	76.4	651	7 CO783493	CO783493 BL018A_F0
C 24	16.8	76.4	651	8 AZ570026	AZ570026 269PVF02

25	16.8	76.4	661	4 BI604551	BI604551 RH69961.5
26	16.8	76.4	665	9 CL672094	CL672094 PR10166b
27	16.8	76.4	670	4 BI230967	BI230967 RE16641.5
28	16.8	76.4	679	4 BI234016	BI234016 RE29651.5
29	16.8	76.4	679	4 BI633310	BI633310 SD27370.5
30	16.8	76.4	683	4 BI228896	BI228896 RE26554.5
31	16.8	76.4	699	2 BF867544	BF867544 963092D11
32	16.8	76.4	729	8 BH101008	BH101008 RPCI-24-3
C 33	16.8	76.4	732	1 AA698937	AA698937 HL05977.3
C 34	16.8	76.4	955	5 BX838580	BX838580 BX838580
C 35	16.8	76.4	2248	3 CNS0A7TY	BX823496 Arabidops
C 36	16.4	74.5	513	5 BQ903224	BQ903224 Ta03_11b0
C 37	16.4	74.5	551	2 BE191646	BE191646 MD0027 Me
C 38	16.4	74.5	555	7 CO751788	CO751788 Mdfct3052
39	16.4	74.5	544	4 BM581827	BM581827 170006872
40	16.4	74.5	549	5 BX100177	BX100177 BX100177
41	16.4	74.5	1101	9 CNS00270	AL097326 Drosophila
C 42	16.4	74.5	1112	2 CNS02PYJ	AL208612 Tetraodon
C 43	16.4	74.5	1180	2 BE548142	BE548142 601073106
C 44	16.2	73.6	182	1 AA084430	AA084430 zf76903.r
45	16.2	73.6	185	7 CFI18380	CFI18380 f8496.r1

ALIGNMENTS

RESULT 1
W66011/c
LOCUS
DEFINITION
TgESTzy7le08.r1 TGME49 Tachyzoite cDNA Toxoplasma gondii cDNA clone
ESTzy7le08.r1 5', mRNA sequence.
W66011
VERSION
W66011.1 GI:1374199
KEYWORDS
EST.
SOURCE
Toxoplasma gondii
ORGANISM
Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
REFERENCE
1 (bases 1 to 297)
AUTHORS
Hehl, A., Manger, I., Marra, M., Sibley, L.D., Ajioke, J.A.,
Aslett, M.A., Dietrich, N., Dubuque, T., Hillier, L., Kucaba, T.,
Wan, K.L., Waterston, R.H. and Boothroyd, J.
TITLE
WashU-Merck-Stanford-NIH Toxoplasma EST project
JOURNAL
Unpublished (1996)
COMMENT
Contact: Marra M
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@watson.wustl.edu
David Sibley at toxoest@borcim.wustl.edu for further information
relating to organism, clone or library availability.
Seq primer: T3
High quality sequence stop: 217.
Location/Qualifiers
1..297
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="ME49, clone PDS"
/db_xref="taxon:5811"
/clone="tgzy7le08.r1"
/lab_host="XLI-Blue MRF"
/notes="Vector: Lambda Zap; Site 1: EcoRI; Site 2: XhoI;
Toxoplasma parasites were grown in human foreskin
fibroblasts cultures in vitro. The library was
constructed by A. Hehl and I. Manger, Stanford University.
cDNAs were synthesized from polyA mRNAs by oligo d(T)
priming and directionally cloned into the EcoRI and XhoI
sites of the Lambda Zap vector using the ZAP-cDNA
synthesis kit (Statagene). Warning: the library contains a
small percentage of human cDNAs derived from the human

FEATURES
source

```
ORIGIN
Query Match      80.9%; Score 17.8; DB 7; Length 297;
Best Local Similarity 90.5%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAAGCGGAAC 22
    |||||
Db 212 CACCTGAACGTTAGGCGGAAC 192

RESULT 2
BJ434099 396 bp mRNA linear EST 13-MAR-2002
LOCUS BJ434099 dictyostelium discoidium cDNA library, VF Dictyostelium
DEFINITION discoidium cDNA clone ddv23k20 3', mRNA sequence.
ACCESSION BJ434099.1 GI:19408821
VERSION BJ434099
KEYWORDS EST.
SOURCE Dictyostelium discoidium
ORGANISM Dictyostelium discoidium
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE Full length cDNA of Dictyostelium discoidium at the vegetative
JOURNAL stage
COMMENT Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gene.nig.ac.jp.

FEATURES
source
1..396
    /organism="Dictyostelium discoidium"
    /mol_type="mRNA"
    /strain="AX4"
    /db_xref="taxon:44689"
    /clone="ddv23k20"
    /sex="mat A"
    /dev_stage="Growth phase"
    /clone_lib="Dictyostelium discoidium cDNA library, VF"

ORIGIN
Query Match      80.9%; Score 17.8; DB 4; Length 396;
Best Local Similarity 90.5%; Pred. No. 3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAAGCGGAA 21
    |||||
Db 211 CCACCTGAACGACCAGCGGAA 231

RESULT 3
BM176732/c 547 bp mRNA linear EST 06-DEC-2001
LOCUS BM176732 TgESTzya98b10.y1 TgrH Tachyzoite Subtracted cDNA Library Toxoplasma
DEFINITION gondii cDNA clone TgESTzya98b10.y1 5', mRNA sequence.
ACCESSION BM176732
VERSION BM176732
KEYWORDS EST.
SOURCE Toxoplasma gondii
ORGANISM Toxoplasma gondii
REFERENCE Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
AUTHORS Sarcocystidae; Toxoplasma.
1 (bases 1 to 547)
Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I.,
Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.

host cells."

TITLE
JOURNAL
COMMENT

Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
Toxoplasma EST Project
Unpublished (2001)
Contact: Clifton, S.

Toxoplasma EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@watson.wustl.edu

Contact David Sibley (toxost@borcim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco
High quality sequence stop: 425.
Location/Qualifiers
1..547
    /organism="Toxoplasma gondii"
    /mol_type="mRNA"
    /strain="RH (Type I)"
    /db_xref="taxon:5811"
    /clone="TgESTzya98b10.y1"
    /dev_stage="Tachyzoite"
    /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
    /clone_lib="TgrH Tachyzoite Subtracted cDNA Library"
    /note="Vector: pBluscript SK-; Site 1: EcoRI; Site 2:
XhoI; Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
originally constructed by K.L.Wan, Cambridge University.
cDNAs were synthesized from polyA RNAs by oligo d(T)
priming and directionally cloned into the EcoRI to XhoI
sites of the Lambda ZapII vector using the Zap-cDNA
synthesis kit (Stratagene). The primary cDNA library was
mass excised as phagemid using ExAssist helper phage
(Stratagene). Phagemid DNA was extracted by
phenol-chloroform method, and hybridized against a pool of
over-represented ESTs (N=12, from 5596 previous reads).
The subtracted library was electroporated into DH10B
(GeneHog, Invitrogen, Inc). WARNING: the library contains
a small percentage of cDNAs derived from the human host
cells. Library Source: David Sibley, Washington
University."
```


Toxoplasma EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxo@watson.wustl.edu
Contact David Sibley (toxosent@borderim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco
High quality sequence stop: 435.

FEATURES

source
1..701
/organism="Toxoplasma gondii"
/mol_type="mRNA"
/strain="RH (Type 1)"
/db_xref="taxon:5811"
/clone="TgESTztd89h12.y1"
/dev_stage="Tachyzoite"
/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
/clone_lib="TgRH Tachyzoite Norm 7 cDNA Library"
/notes="Vector: pBluscript SK; Site 1: EcoRI; Site 2:
XhoI; Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
originally constructed by K.L.Wan, Cambridge University.
cDNAs were synthesized from polyA RNAs by oligo d(T)
priming and directionally cloned into the EcoRI to XhoI
sites of the Lambda ZapII vector using the ZAP-cDNA
synthesis kit (Stratagene). The primary cDNA library was
mass excised as phagemid using ExAssist helper phage
(Stratagene). Phagemid DNA was extracted by
phenol-chloroform method, and hybridized against a pool of
highly abundant genes which were derived from short-cycle
PCR of the primary cDNA library. The normalized library
was electroporated into DH10B (GeneHog, Invitrogen, Inc).
WARNING: the library contains a small percentage of cDNAs
derived from the human host cells."

ORIGIN

Query Match 80.9%; Score 17.8; DB 6; Length 701;
Best Local Similarity 90.5%; Pred. No. 3.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAGCGGAC 22
|||||
Db 374 CACCTGAACGATAGCGGAC 354

RESULT 5

CC180655/c 102 bp DNA linear GSS 05-MAY-2003
LOCUS 01S0568-02A1-B03 UniformMu MutAIL Library zea mays genomic clone
DEFINITION 01S0568-02A1-B03, genomic survey sequence.

ACCESSION CC180655
VERSION CC180655.1 GI:30352697
KEYWORDS GSS.

SOURCE

Zeae mays
Zeae mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 102)
Latshaw,S., Tan,B.-C., Settles,A.M. and McCarty,D.R.
Sequence tagged transposon insertions from the UniformMu maize
population

AUTHORS

Unpublished (2003)
Contact: Donald R. McCarty

JOURNAL

Plant Molecular and Cellular Biology Program
University of Florida
PO 110690 Gainesville, FL 32611-0690, USA

TITLE

Tel: 352-392-1928 x322
Email: drmc@ufl.edu
Sequence flanking probable Mu insertion site in UniformMu line:
01S0568-02, Primer set: A

FEATURES

source
Location/Qualifiers

1..102
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="W22 (ACR, bz1-m9)"
/cultivar="UniformMu"
/db_xref="taxon:4577"
/clone_lib="01S0568-02A1-B03"
/note="Vector: TOPO-PCR4; DNA flanking Mu transposon
insertions in Mu inactive lines were extracted from the
UniformMu maize population by the thermo asymmetric
interlaced PCR (TAIL) protocol using primers specific for
the Mu terminal inverted repeat and a set of 16 arbitrary
primers. Amplicons were size enriched using Sepharose 400
spin columns and cloned into the TOPO PCR4 vector."

ORIGIN

Query Match 79.1%; Score 17.4; DB 8; Length 102;
Best Local Similarity 94.7%; Pred. No. 4.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCACCTGACGATAGCGG 19
|||||

Db 51 CCACCTGACGATAGCGG 33
|||||

RESULT 6

BM774282 529 bp mRNA linear EST 25-APR-2002
ra91e12.y1 Meloidogyne incognita egg pAMP1 Bird Meloidogyne
incognita CDNA 5' similar to TR:Q09165 Q09165 HYPOTHEICAL 1368.6
KD PROTEIN K07E12.1 IN CHROMOSOME III. ; mRNA sequence.

ACCESSION

BM774282
VERSION BM774282.1 GI:19103896

KEYWORDS

EST
Meloidogyne incognita (southern root-knot nematode)

SOURCE

Meloidogyne incognita
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Meloidogyminae; Meloidogyne.

REFERENCE

1 (bases 1 to 529)
McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
Wyllie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
Tsagarisvilli,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)

TITLE

Contact: McCarter JP

JOURNAL

The Washington Univ. Nematode EST Project, 1999

COMMENT

Washington Univ. Nematode EST Project, 1999
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

The library was constructed by David Bird (david.bird@cnsu.edu).

DNA Sequencing by: Washington University Genome Sequencing Center

High quality sequence stop: 401.

FEATURES

source
Location/Qualifiers

1..529
/organism="Meloidogyne incognita"
/mol_type="mRNA"
/db_xref="taxon:6306"
/dev_stage="egg"
/lab_host="DH10B"
/clone_lib="Meloidogyne incognita egg pAMP1 Bird"
/note="Vector: pAMP1 (Gibco); Site 1: NotI; Site 2: SalI;
The library was provided by Dr. David Bird at North
Carolina State University, Raleigh, NC
(david_bird@cnsu.edu). The cDNA was made by using Dynabead

oligo-dT priming (Dyna). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMP1."

ORIGIN

Query Match 79.1%; Score 17.4; DB 4; Length 529;
Best Local Similarity 94.7%; Pred. No. 5e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ACCTGAACGATAGCGGAA 21
|||||
Db 43 ACCTGAACGATAGCGGAA 61
|||||

RESULT 7

CG754242/c

LOCUS 1969 bp DNA linear GSS 24-OCT-2003
DEFINITION P049-3-E01.za Ppa EcORI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.

ACCESSION

CG754242

CG754242.1 GI:37979536

KEYWORDS

GSS

SOURCE

Pristionchus pacificus

ORGANISM

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.

REFERENCE

1 (bases 1 to 1969)

AUTHORS

Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
Buntjer,J., van der Meulen,M. and Sommer,R.J.

TITLE

An integrated physical and genetic map of the nematode Pristionchus
pacificus

JOURNAL

Mol. Genet. Genomics 269 (5), 715-722 (2003)

MEDLINE

22835951

PUBMED

12884007

COMMENT

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spenannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: ralf.sommer@tuebingen.mpg.de

Class: BAC ends.

FEATURES

Location/Qualifiers

1..1969

/organism="Pristionchus pacificus"

/mol_type="genomic DNA"

/strain="California"

/db_xref="taxon:54126"

/clone_lib="Ppa EcORI BAC Library"

/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

Query Match 79.1%; Score 17.4; DB 9; Length 1969;
Best Local Similarity 94.7%; Pred. No. 5.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCACCTGAACGATAGCGG 19
|||||
Db 1166 CCACCTGAACGATAGCGG 1148
|||||

RESULT 8

BB378617/c

LOCUS

BB378617

DEFINITION

329 bp mRNA linear EST 13-JUL-2000

musculus cDNA clone C130095A16 3' similar to X57337 M.musculus

CDNAp14 mRNA, mRNA sequence.

ACCESSION

BB378617

VERSION

BB378617.1 GI:9091490

KEYWORDS

EST.

SOURCE
ORGANISM

Mus musculus (house mouse)

REFERENCE

1 (bases 1 to 329)

AUTHORS

Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T.,
Carninci,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,
Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,
Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,
Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C.,
Kusabe,M., Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H.,
Okazaki,Y., Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K.,
Shibata,K., Shibata,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Sugahara,Y., Suzuki,H., Suzuki,H., Tagawa,A.,
Takahashi,P., Tominaga,N., Tova,T., Tsunoda,Y., Watahiki,A.,
Watanabe,S., Yamamura,T., Yamanaka,I., Yano,R., Yasunishi,A.,
Yokota,T., Yoshida,K., Yoshiki,A., Yoshino,M., Muramatsu,M. and
Hayashizaki,Y.

TITLE

RIKEN Mouse ESTs (Konno,H., et al.)

JOURNAL

Unpublished (2000)

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S.,
Sasaki,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Thermotabilization and thermoactivation of thermolabile enzymes by
trehalose and its application for the synthesis of full length
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh,M., Kitanai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.

FEATURES

Location/Qualifiers

1..329

/organism="Mus musculus"

/mol_type="mRNA"

/strains="CS7BL/6J"

/db_xref="taxon:10090"

/clone="C130095A16"

/sex="mixed"

/tissue_type="head"

/dev_stage="16 days embryo"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 16 days embryo
head"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTTCGAGTTAATTAATTAATTCCTCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. a
modified pluescript KS(+) after bulk excision from Lambda
FLC I"

ORIGIN

```

Query Match      78.2%; Score 17.2; DB 2; Length 329;
Best Local Similarity 86.4%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
    |||||
Db 148 CCACCTGACCATCAGCGAAC 127

RESULT 9
AOL173857      417 bp DNA linear GSS 17-OCT-1998
LOCUS HS_3199_B1_F04_MR_CIT Approved Human Genomic Sperm Library D Homo
DEFINITION sapiens genomic clone Plate=3199 Col=7 Row=L, Genomic survey
sequence.
ACCESSION AQL173857
VERSION AQL173857.1 GI:3571224
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 417)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,K., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3199 row: L column: 7
Class: BAC ends
High quality sequence stop: 417.
FEATURES
source
    Location/Qualifiers
    1..417
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clone="Plate=3199 Col=7 Row=L"
        /sex="male"
        /clone_lib="CIT Approved Human Genomic Sperm Library D"
        /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
        E-Coli DH10B"

ORIGIN
Query Match      78.2%; Score 17.2; DB 8; Length 417;
Best Local Similarity 86.4%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
    |||||
Db 99 CCACCTGAGCGATAGCTGAAC 120

RESULT 10
CA067617/c
LOCUS CA067617      582 bp mRNA linear EST 23-SEP-2003
DEFINITION SCQSDA1055C02.b AD1 Saccharum officinarum cDNA clone SCQSDA1055C02
3', mRNA sequence.
ACCESSION CA067617
VERSION CA067617.1 GI:34919141
KEYWORDS EST.
SOURCE Saccharum officinarum
    Location/Qualifiers
    1..582
        /organism="Saccharum officinarum"
        /mol_type="mRNA"
        /db_xref="taxon:4547"
        /clone="SCQSDA1055C02"
        /lab_host="DH10B"
        /clone_lib="AD1"
        /note="Organ: seedlings inoculated with Gluconacetobacter
        diazotrophicans; Vector: pSport1; Site:1. SalI; Site_2:
        NotI; An unidirectional cDNA library generated from
        seedlings inoculated with Gluconacetobacter
        diazotrophicans]. cDNA was prepared from polyA+ mRNA using
        SuperScript Plasmid System Kit (Invitrogen). The
        double-strand cDNAs were fractionated in a sepharose
        CL-2B 40cm-columns and fragments sizing between 0.8 and
        1.5 Kb were directionally cloned into the vector. Details
        of each source of RNA and library construction can be
        obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN
Query Match      78.2%; Score 17.2; DB 6; Length 582;
Best Local Similarity 86.4%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22
    |||||
Db 434 CCACCTGACGACGAGACTGAC 413

RESULT 11
AU022955
LOCUS AU022955      587 bp mRNA linear EST 20-OCT-1998
DEFINITION AU022955 Mouse unfertilized egg cDNA Mus musculus cDNA clone
J0423C03 3', mRNA sequence.
ACCESSION AU022955
VERSION AU022955.1 GI:3393302
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 587)
AUTHORS Ko,M.S.H, Kitchen,J.R., Wang,X., Threat,T.A., Sun,T., DePalma,G.E.,
Liang,Y., Kargul,G.J., Sharara,R. and Doi,H.
TITLE Systematic analyses of genes expressed in unfertilized mouse eggs
(the ERAIO/Doi Project at Wayne State University) (Ko,M.S.H. et
al.)
JOURNAL Unpublished (1998)
COMMENT Contact: Hirofumi Doi
Doi Biosymetry Project, ERATO
Japan Science and Technology Corporation (JST)

```

Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.

1 (bases 1 to 582)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

CONTACT: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089

Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br

Plate: 055 row: C column: 02
Seq primer: SP6 Promoter primer.

FEATURES
source

Location/Qualifiers
1..582
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCQSDA1055C02"
/lab_host="DH10B"
/clone_lib="AD1"

/note="Organ: seedlings inoculated with Gluconacetobacter
diazotrophicans; Vector: pSport1; Site:1. SalI; Site_2:
NotI; An unidirectional cDNA library generated from
seedlings inoculated with Gluconacetobacter
diazotrophicans]. cDNA was prepared from polyA+ mRNA using
SuperScript Plasmid System Kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CL-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 78.2%; Score 17.2; DB 6; Length 582;
Best Local Similarity 86.4%; Pred. No. 6.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22

|||||
Db 434 CCACCTGACGACGAGACTGAC 413

/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
E-Coli DH10B"

ORIGIN

Query Match 78.2%; Score 17.2; DB 8; Length 417;
Best Local Similarity 86.4%; Pred. No. 6.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCGGAAC 22

|||||
Db 99 CCACCTGAGCGATAGCTGAAC 120

RESULT 10

CA067617/c
LOCUS CA067617 582 bp mRNA linear EST 23-SEP-2003
DEFINITION SCQSDA1055C02.b AD1 Saccharum officinarum cDNA clone SCQSDA1055C02
3', mRNA sequence.

ACCESSION CA067617

VERSION CA067617.1 GI:34919141

KEYWORDS EST.

SOURCE Saccharum officinarum

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome
JOURNAL Unpublished (1998)
COMMENT Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: GGAAACAGCTATGACCATG
 Class: BAC ends
 High quality sequence stop: 408.
FEATURES Location/Qualifiers
 source
 1..615
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0020H02r"
 /tissue_type="protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII; Rice blast is one of the most devastating fungal diseases of rice world wide. It is a filamentous ascomycete with a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

FEATURES

source

Query Match 77.3%; Score 17; DB 8; Length 615;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 3 ACCTGAACGATAGCGG 19
 |||||
Db 493 ACCTGAACGATAGCGG 477

RESULT 15
CNS014C4
LOCUS 1038 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC BACN1105 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION AL103966
VERSION AL103966.1 GI:5615577
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1038)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - <http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOBAC11.

FEATURES

source

Location/Qualifiers
 1..1038
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACN11L05"
 /clone_lib="DrosBAC"
 /plasmid="pBelOBAC11"
 /note="end : SP6"

ORIGIN

Query Match 77.3%; Score 17; DB 9; Length 1038;
Best Local Similarity 85.7%; Pred. No. 8.6e+02;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATAGCGGAAC 22

Db 36 CACCTGAACGACAGCGGAH 56

RESULT 16

BF742895
LOCUS IL2-BT0734-041000-178-A04 BT0734 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF742895
ACCESSION BF742895
VERSION BF742895.1 GI:12069571
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 407)
 Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,P., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&t2=IL2-BT0734-041000-178-A04&t3=2000-10-04&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 406.
Location/Qualifiers
 1..407
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BT0734"

QY 2 CACCTGAACGATAGCGGAAC 22

Db 36 CACCTGAACGACAGCGGAH 56

RESULT 16

BF742895

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&t2=IL2-BT0734-041000-178-A04&t3=2000-10-04&t4=1>)

Seq primer: puc 18 forward

High quality sequence stop: 406.

Location/Qualifiers

1..407

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BT0734"

/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES

source

Query Match 77.3%; Score 17; DB 9; Length 1038;
Best Local Similarity 85.7%; Pred. No. 8.6e+02;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

ORIGIN

```

Query Match      76.4%; Score 16.8; DB 2; Length 407;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CACCTGAACGATTAAGCGGA 21
Db 357 CAGCTGACCGATAAGCGGA 376

RESULT 17
BZ291938/c
LOCUS
DEFINITION
SALK_122722.53.75.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_122722.53.75.x, genomic
survey sequence.
ACCESSION
BZ291938 GI:24338224
VERSION
BZ291938.1
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 425)
AUTHORS
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL
Unpublished (2001)
COMMENT
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated exon of At3g26600.
Class: TDNA tagged.
FEATURES
Location/Qualifiers
1..425
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone_lib="SALK_122722.53.75.x"
/clone="SALK_122722.53.75.x"
/note="PCR was performed on Arabidopsis thaliana TDNA insertion lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
source
Query Match      76.4%; Score 16.8; DB 8; Length 425;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATTAAGCGGAAC 22
Db 49 ACCTGAACAAATAATCGGAAC 30

RESULT 18
CR341213/c
LOCUS
DEFINITION
Medicago truncatula BAC ends cultivar Jemalong A17 of Medicago
truncatula, genomic survey sequence.
ACCESSION
CR341213 GI:44911548
VERSION
CR341213.1
KEYWORDS
GSS.

SOURCE
Medicago truncatula (barrel medic)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 578)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (25-FEB-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
FEATURES
Location/Qualifiers
1..578
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="Jemalong A17"
/db_xref="taxon:3880"
/clone_lib="MTE1"
/note="Vector: pIndigoBAC ; Site 1: EcoRI ; Site 2: EcoRI
; Debelle F. and Chalhoub B.-Genoscope sequence ID :
mtel-73M16FM1"

ORIGIN
source
Query Match      76.4%; Score 16.8; DB 9; Length 578;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATTAAGCGGAAC 22
Db 343 ACCTGAACAAATAAGCGGTAC 324

RESULT 19
A1389889
LOCUS
DEFINITION
GH21181.5prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH21181 5prime similar to M26400:
&gt;-Spec Fgmn0003470 PID:g158489 SWISS-PROT:P13395, mRNA
sequence.
ACCESSION
A1389889
VERSION
A1389889.1 GI:4203900
KEYWORDS
EST.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 634)
AUTHORS
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
TITLE
BDGP/HMI Drosophila EST Project
JOURNAL
Unpublished (2001)
COMMENT
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic sequence AC005557
Plate: 211 row: G column: 9
High quality sequence stop: 416.
Location/Qualifiers
1..634
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH21181"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/clone_lib="GH Drosophila melanogaster head pOT2"
/note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:

```

XhoI; Sized fractionated cDNAs were directly ligated into pOT2. Plasmid cDNA library."

ORIGIN

Query Match 76.4%; Score 16.8; DB 1; Length 634;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCGGAAC 22

||||| ||||| ||||| |||||

Db 59 ACCTGAACGAGAGTGGAC 78

RESULT 20

BI363699

LOCUS

DEFINITION RE48622.5prime RE Drosophila melanogaster normalized Embryo pFLC-1
Drosophila melanogaster cDNA clone RE48622 5 similar to smc:
FBan0011561 'G protein linked receptor' located on: 2L 21B5-21B5;;
05/13/2001, mRNA sequence.

ACCESSION

VERSION BI363699

KEYWORDS GI:15059727

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,
Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E.,
George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G.,
Miser, S., Mungall, C. J., Nunoo, J., Pacieb, J., Paragas, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and
Rubin, G. M.

TITLE BDGP/HMI RE Drosophila EST Project

JOURNAL Unpublished (2001)

COMMENT Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
hit genomic AE003590: arm:2L [1,301639] estimated-cyto:40P7-21B7:
05/13/2001

Plate: RE.486 row: B column: 10

High quality sequence stop: 502.

FEATURES

source

Location/Qualifiers

1. .637

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="RE48622"

/sex="male and female"

/dev_stage="0-24 hours mixed stage embryonic"

/lab_host="DH5-alpha Tona"

/clone_lib="RE Drosophila melanogaster normalized Embryo

pFLC-1"

/note="Organ: embryo; Vector: pFLC1; Site.1: XhoI; Site.2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 76.4%; Score 16.8; DB 4; Length 637;

Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCGGAAC 22

||||| ||||| ||||| |||||

Db 603 ACCTGAACGATAGCTGAAC 622

RESULT 21

AA979615

LOCUS

DEFINITION

649 bp mRNA linear EST 02-DEC-2003
LD34308.5prime LD Drosophila melanogaster embryo pOT2 Drosophila
melanogaster cDNA clone LD34308 5 similar to smc: FBan0011561
GO: [patched receptor ligand receptor (GO:0008158); smo receptor
signalling pathway (GO:0007224); G protein linked receptor
(GO:0004930); plasma membrane (GO:0005886); plasma membrane (GO:0004930); plasma membrane (GO:0005886); mRNA sequence.
[description truncated]: 08/13/2002, mRNA sequence.

ACCESSION

VERSION AA979615

KEYWORDS GI:13769047

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
On May 26, 1998 this sequence version replaced gi:3157020.
Other ESTs: LD34308.3prime
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
hit genomic AE003590: arm:2L [1,301639] estimated-cyto:40P7-21B7:
04/10/2001

Plate: LD.343 row: A column: 8

High quality sequence stop: 649

POLYA=No.

FEATURES

source

Location/Qualifiers

1. .649

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="LD34308"

/sex="male and female"

/dev_stage="0 to 24 hours mixed stage embryonic"

/lab_host="Xhl Blue"

/clone_lib="LD Drosophila melanogaster embryo pOT2"
/note="Organ: embryo; Vector: pOT2; Site.1: EcoRI; Site.2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2."

ORIGIN

Query Match

Best Local Similarity 76.4%; Score 16.8; DB 1; Length 649;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCGGAAC 22

||||| ||||| ||||| |||||

Db 581 AGCTGAACGATAGCTGAAC 600

RESULT 22

BI229963

LOCUS

DEFINITION

651 bp mRNA linear EST 11-JUL-2001
RE27868.5prime RE Drosophila melanogaster normalized Embryo pFLC-1
Drosophila melanogaster cDNA clone RE27868 5 similar to smc:
FBan0011561 'G protein linked receptor' located on: 2L 21B5-21B5;;
04/12/2001, mRNA sequence.

ACCESSION

VERSION BI229963

KEYWORDS GI:14697227

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

cushion. Purified DNA was digested with mung bean nuclease in the presence of 44% formamide at 500C as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1989. Nucleic Acids Research 16:6883-6896). Digested DNA was blunt-ended using T4 DNA Polymerase and size fractionated over a Sepharose CL-2B column. Fractions in the size range 500bp-4kb were ligated into the Eco RV site of plusescript SK(+), and E. coli XL-10 Gold transformed with the ligation mixture."

ORIGIN

Query Match 76.4%; Score 16.8; DB 8; Length 651;
Best Local Similarity 90.0%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

QY 1 CCACCTGAACGTAAGCGGA 20
|||||
Db 262 CCACCTGAACGTAAGCGGA 243

RESULT 25

BI604551

LOCUS

DEFINITION RH69961.5prime RH Drosophila melanogaster normalized Head pFlc-1
Drosophila melanogaster cDNA clone RH69961 5 similar to smc;
FBan011561 GO: [patched receptor ligand receptor (GO:0008158); smc
receptor signalling pathway (GO:007224); G protein linked receptor
(GO:0004930); plasma membrane (GO:0005886); plasma membrane
(GO:0005886)]; plasma membrane (GO:0, mRNA sequence.

ACCESSION

BI604551

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Stapleton, M., Brokstein, P., Hong, L., Tyler, D., Berman, B.,
Carlson, J., Champe, M., Chavez, C., Dorsett, V., Farfan, D., Friese, E.,
George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G.,
Misra, S., Mungall, C. J., Nuncio, J., Pacleb, J., Paragas, V., Park, S.,
Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Ceiniker, S. and
Rubin, G. M.

TITLE

JOURNAL

COMMENT

BDGP/HMI RH Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu

hit genomic AE003590: arm:2L [1,301639] estimated-cyto:407-21B7:
08/24/2001

Plate: RH.699 row: F column: 1

High quality sequence stop: 516.

FEATURES

source

1. .661

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="RH69961"

/sex="male and female"

/dev_stage="Adult"

/lab_host="DH5-alpha Tona"

/clone_lib="RH Drosophila melanogaster normalized Head

pFlc-1"

/note="Organ: head; Vector: pFlc1; Site: 1: XhoI; Site 2:

BamHI; Library was kindly generated by Piero Carninci at

the RIKEN. The library was normalized and excised using

Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 76.4%; Score 16.8; DB 4; Length 661;
Best Local Similarity 90.0%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

QY 3 ACCTGAACGTAAGCGGAAC 22

|||||

Db 604 ACCTGAACGTAAGCTGAAC 623

RESULT 26

CL672094

LOCUS

DEFINITION

PR10166b_H11 - PR10166b.B21 (665) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.

CL672094

CL672094.1 GI:50171951

GSS.

Pristionchus pacificus

Pristionchus pacificus

Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;

Neodiplogasteridae; Pristionchus.

1 (bases 1 to 665)

Srinivasan, J., Octo, G. W., Kahlow, U., Geisler, R. and Sommer, R. J.

AppaDB: an Acdb database for the nematode satellite organism

Pristionchus pacificus

Nucleic Acids Res. 32 (1), D421-D422 (2004)

Contact: Sommer RJ

Evolutionary Biology

Max-Planck-Institute for Developmental Biology

Spemannstr. 37-39, Tuebingen D-72076, Germany

Tel: 00497071601371

Fax: 00497071601498

Email: raif.sommer@tuebingen.mpg.de

This library was generated at Caltech, Pasadena, USA and end

sequenced at Vancouver, Canada.

Seq primer: T7

Class: fosmid ends.

Location/Qualifiers

1. .665

/organism="Pristionchus pacificus"

/mol_type="genomic DNA"

/strain="California"

/db_xref="taxon:54126"

/clone_lib="Mixed stage fosmid library of P. pacificus

var. California"

/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match

Best Local Similarity

Matches

18;

Conservative

0;

Mismatches

2;

Indels

0;

Gaps

0;

QY

2

CACCTGAACGTAAGCGGAA 21

|||||

Db

143

CACCTCAACGTAATCGGAA 162

RESULT 27

BI230967

LOCUS

DEFINITION

REI16641.5prime RE Drosophila melanogaster normalized Embryo pFlc-1

Drosophila melanogaster cDNA clone REI16641 5 similar to smc;

FBan011561 'G protein linked receptor' located on: 2L 21B5-21B5;;

05/11/2001, mRNA sequence.

ACCESSION

BI230967

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

```

REFERENCE
AUTHORS
1 (bases 1 to 670)
Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B.,
Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E.,
George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G.,
Misra,S., Mungall,C.J., Nunoo,J., Pacieb,J., Paragas,V., Park,S.,
Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Ceiniker,S. and
Rubin,G.M.
BDGP/HMI RE Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003590: arm:2L [1,301639] estimated-cyto:40F7-21B7:
05/11/2001
Plate: RE.166 row: D column: 5
High quality sequence stop: 540.
Location/Qualifiers
1..670
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE16641"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/notes="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."

FEATURES
source
1..670
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="RE29651"
/sex="male and female"
/dev_stage="0-24 hours mixed stage embryonic"
/lab_host="DH5-alpha Tona"
/clone_lib="RE Drosophila melanogaster normalized Embryo
pFlc-1"
/notes="Organ: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
BamHI; Library was kindly generated by Piero Carninci at
the RIKEN. The library was normalized and excised using
Cre recombinase. Plasmid cDNA library."

ORIGIN
Query Match 76.4%; Score 16.8; DB 4; Length 670;
Best Local Similarity 90.0%; Pred.No.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCGGAC 22
| | | | | | | | | | | | | | | | | |
Db 603 AGCTGAACGATAGCTGAAC 622

RESULT 29
BI633310
LOCUS
DEFINITION
BI633310 679 bp mRNA linear EST 10-SEP-2001
SDT2 Drosophila melanogaster Schneider L2 cell culture
PBan0011561 GO: [patched receptor ligand receptor (GO:0008158); smo
receptor signalling pathway (GO:0007224); G protein linked receptor
(GO:0004930); plasma membrane (GO:0005886); plasma membrane
(GO:0005886); plasma membrane (GO:00, mRNA sequence.

ACCESSION
VERSION
BI633310.1 GI:15535520
KEYWORDS
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 679)
AUTHORS
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003590: arm:2L [1,301639] estimated-cyto:40F7-21B7:
05/22/2001
Plate: SD.273 row: F column: 10
High quality sequence stop: 632.
Location/Qualifiers
1..679
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="SD27370"
/lab_host="DH5-alpha"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture pOT2"

FEATURES
source
1..679
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="SD27370"
/lab_host="DH5-alpha"
/clone_lib="SD Drosophila melanogaster Schneider L2 cell
culture pOT2"

```


AUTHORS

Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M., Tsegaye, G., Geer, K., Kroi, M., Shvartsbeyn, A., Gebregorgis, E., Russell, D., de Jong, P., and Fraser, C.M.
 Mouse BAC End Sequences from Library RPCI-24
 Unpublished (1999)

JOURNAL
COMMENT

Other_GSSB: RPCI-24-378D7.TV
 Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 378 row: D column: 7
 Seq primer: SP6
 Clones: BAC ends.

FEATURES

source

Location/Qualifiers

```

1..729
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-378D7"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site 1: BamH1; Site 2: BamH1;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."

```

ORIGIN

```

Query Match      76.4%; Score 16.8; DB 8; Length 729;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 3 ACCTGAACGATAGCGGAC 22

```

|||||
Db 359 ACCTGAACGATAGCAAAAC 340

```

RESULT 33

AA698937/c

LOCUS

```

DEFINITION      AA698937 732 bp mRNA linear EST 19-APR-2001
                  HL05977.3prime HL Drosophila melanogaster head BlueScript
                  Drosophila melanogaster cDNA clone HL05977 3prime similar to
                  M26400: aagr;-Spec F8gm0003470 PID:g158489 SWISS-PROT:P13395, mRNA
                  sequence.

```

ACCESSION

AA698937

AA698937.1 GI:2701866

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 732)

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

Lewis, S. and Rubin, G.M.

BDGP/HMI Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu

Based upon the presence of a XhoI site followed by a run of 14 or more T residues at the beginning of the sequence, this clone probably contains an inverted insert. The sequence has been trimmed and the T residues removed.

Plate: 59 row: G column: 5

High quality sequence stop: 512.

FEATURES

source

Location/Qualifiers

```

1..732
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="HL05977"
/sex="male and female"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="HL Drosophila melanogaster head BlueScript"
/note="Organ: head--brain & sensory organ; Vector:
BlueScript SK; Site 1: EcoRI; Site 2: XhoI; Constructed
using Stratagene ZAP-cDNA Synthesis kit. Oligo dT-primed
and directionally cloned at EcoRI and XhoI in BlueScript
SK(+/-)"

```

ORIGIN

```

Query Match      76.4%; Score 16.8; DB 1; Length 732;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 3 ACCTGAACGATAGCGGAC 22

```

|||||
Db 129 ACCTGAACGAGAAAGTGAAC 110

```

RESULT 34

BX838580/c

LOCUS

```

DEFINITION      BX838580 955 bp mRNA linear EST 11-FEB-2004
                  Arabidopsis thaliana Flowers and buds Col-0 Arabidopsis
                  thaliana cDNA clone GSLTFFB732A02 SPRIM, mRNA sequence.

```

ACCESSION

BX838580

BX838580.1 GI:42532663

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana

Arabidopsis thaliana

Eukaryota; Viridiplantae;

Spermatophyta; Magnoliophyta;

rosids; eurosids II; Brassicales;

1 (bases 1 to 955)

Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C.,

Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V.,

Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:

A Combined Approach to Evaluate and Improve Arabidopsis Genome

Annotation

Unpublished (2004)

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

The sequences are based on single pass reads.

Life technologies (a division of invitrogen) members carried out

full-length libraries construction: Temple G.

Genoscope members carried out sequencing and annotation: Castelli

V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,

Schachter V., Weissenbach J., Salanoubat M.

URV INRA - Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis

genome released by MIPS (Munich Information center for Protein

Sequences).

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/EST<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

Location/Qualifiers

1..955

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

```

/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="GSLTfB732A02"
/tissue_type="Flowers and buds"
/clone_lib="Arabidopsis thaliana Flowers and buds Col-0"

ORIGIN

Query Match      76.4%; Score 16.8; DB 5; Length 955;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCGGAAC 22
    ||||| ||||| ||||| |||||
Db 822 ACCTGAACAATAATCGGAAC 803

RESULT 35
CNS0A7TY/c
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTfS442D07 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX823496.1 GI:42466469
VERSION
HTC; GSUT cDNA.
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 2248)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J., and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2248)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
            source
            Location/Qualifiers
                1..2248
                /organism="Arabidopsis thaliana"
                /mol_type="mrna"
                /strain="Col-0"
                /db_xref="taxon:3702"
                /clone="GSLTfS442D07"
                /tissue_type="Adult vegetative tissue"
                /plasmid="pCMVSPORT_6"
                1..2248
                /gene="At3g26600"

gene

ORIGIN

Query Match      76.4%; Score 16.8; DB 3; Length 2248;
Best Local Similarity 90.0%; Pred. No. 1.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 ACCTGAACGATAGCGGAAC 22
    ||||| ||||| ||||| |||||
Db 882 ACCTGAACAATAATCGGAAC 863

RESULT 36
BQ903224/c
LOCUS
DEFINITION
Ta03_11b09_R
Ta03_AAFRC_ECORC_Fusarium graminearum inoculated wheat heads
Triticum aestivum cDNA clone Ta03_11b09, mRNA sequence.
ACCESSION
BQ903224
VERSION
BQ903224.1 GI:22302008
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poaceae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 513)
Ouellet,T., Dan,H., Koul,A., Pandeya,R., Chapados,J., Couroux,P.,
De Moors,A., Harris,L.J., Hattori,J.I., Robert,L.S., Singh,J.A.,
Spott,D. and Tinker,N.A.
Expressed Sequence Tags from Wheat Heads 24 Hours after Spray
Inoculation with Fusarium graminearum (part 3)
Unpublished (2002)
JOURNAL
Contact: Ouellet, Therese
Eastern Cereal and Oilseed Research Centre
Agriculture and Agri-food Canada
Nearby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
CANADA
Tel: (613) 759-1658
Fax: (613) 759-1701
Email: ouelletr@agr.gc.ca.

FEATURES
            source
            Location/Qualifiers
                1..513
                /organism="Triticum aestivum"
                /mol_type="mrna"
                /cultivar="FHB148"
                /db_xref="taxon:4565"
                /clone="Ta03_11b09"
                /tissue_type="heads"
                /dev_stage="anthesis"
                /clone_lib="Ta03_AAFRC_ECORC_Fusarium graminearum_inoculate
                d wheat heads"
                /note="Vector: pGEM-T easy; Site 1: EcoRI; Site 2: EcoRI;
                Controlled chamber-grown wheat heads were spray inoculated
                at mid-anthesis with a Fusarium graminearum macroconidial
                suspension (50,000 spores/ml) and kept under intermittent
                misting for 24 hours, then collected and immediately
                frozen in liquid nitrogen."

ORIGIN

Query Match      74.5%; Score 16.4; DB 5; Length 513;
Best Local Similarity 85.0%; Pred. No. 1.6e+03;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 ACCTGAACGATAGCGGAAC 22
    ||||| ||||| ||||| |||||
Db 51 ASCTAAACGATGAGCGGAAC 32

RESULT 37
BE191646/c
LOCUS
DEFINITION
BE191646
Meloiodogyne incognita J2 (#MD99-1) Meloiodogyne incognita
cDNA clone 2882 5' similar to hypothetical protein F34609, mRNA
sequence.
ACCESSION
BE191646
VERSION
BE191646.1 GI:8672030
KEYWORDS
EST.
SOURCE
Meloiodogyne incognita (southern root-knot nematode)

```

ORGANISM	Meloideoglyne incognita Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina; Tylenchoidea; Heteroderidae; Meloideogyninae; Meloideogyne.
REFERENCE	1 (bases 1 to 551)
AUTHORS	Dautova M., Gommers, F. J., Bakker, J. and Smant, G.
TITLE	5' end expressed sequence tags from Meloideoglyne incognita parasitic J2 cDNA library
JOURNAL	Unpublished (2000)
COMMENT	Contact: Smant G / Dautova M Laboratory of Nematology Wageningen University and Research Center Box 8123, Binnenhaven 10, 6709 PD Wageningen, The Netherlands Tel: 31 317 485 254 Fax: 31 317 484 254 Email: Geert.Smant@medew.nema.wau.nl, Wakendona.Dautova@medew.nema.wau.nl Insert Length: 551 Std Error: 0.00 Seq primer: T7 promoter primer High quality sequence stop: 551.

```

FEATURES
source
1. .551
Location/Qualifiers
/organism="Meloiodogyne incognita"
/mol_type="mRNA"
/db_xref="taxon:6306"
/clone="2882"
/dev_stage="second stage preparasitic juveniles (J2)"
/clone_lib="Meloiodogyne incognita J2 (#MD99-1)"
/notes="Vector: pMAK1; Site 1: Sfi IA; Site 2: Sfi IB; cDNA
was synthesized using SMART III oligo, CDS III oligo
d(T)30 (Clontech) and Superscript II reverse
transcriptase (Life Technology). cDNA clones were size
fractionated and directionally ligated in the Sfi IA
restriction site at 5' end and Sfi IB at 3' end of pMAK1.
pMAK1 was derived from the plasmid pCDNA II (Invitrogen)."
```

RESULT 38
 LOCUS C0751788/c
 DEFINITION
 C0751788 linear EST 02-AUG-2004
 Mdftr3052a14.y1 Mdftr Malus x domestica cDNA clone Mdftr3052a14 5' similar to TR:064548 O64548 YUP8H12R.38 PROTEIN. i, mRNA sequence.
 ACCESSION
 C0751788
 VERSION
 C0751788.1 GI:50887042
 KEYWORDS
 EST.
 SOURCE
 Malus x domestica (cultivated apple)
 ORGANISM
 Malus x domestica
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid1; Rosales; Rosaceae; Maloideae; Malus.
 REFERENCE
 1 (bases 1 to 555)
 Korban,S., Vodkin,L., Liu,L., Gasic,K., Gonzales,O., Hernandez,A., Aldwinckle,H., Malnoy,M., Carroll,N., Goldsbrough,P., Orvis,K., Clifton,S., Pape,D., Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Ronko,I., Tsagaris,I., Kennedy,S., Waterston,R. and Wilson,R.
 Apple Functional Genomics grant - NSF 0321702
 TITLE
 Apple Functional Genomics grant - NSF 0321702
 JOURNAL
 Contact: Schuyler S. Korban
 COMMENT
 Apple Functional Genomics grant - NSF 0321702
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Library materials provided by: Schuyler S. Korban Library
constructed by: K. Gasic Library sequenced by: Washington
University Genome Sequencing Center
WashU EST name: asj74a07.Y1
Seq primer: -40UP from Gibco
High quality sequence stop: 546.
Location/Qualifiers

1. .555

/organism="Malus x domestica"
/mol_type="mRNA"
/db_xref="taxon:3750"
/clone="Wdfrt3052a14"
/lab_host="DH10B ampicillin resistant"
/clone_lib="Wdfrt"
/notes="Vector: pBluescript II SK (+); Site 1: NotI;
Site 2: EcoRI; Total RNA was extracted separately from
each stage [young fruitlet (<1cm), young fruitlet (1 cm
dia.), young fruitlet (12cm dia.), maturing fruit 1,
maturing fruit II, mature fruit], using the 'pine tree'
method. Poly(A)+mRNA was isolated twice from total RNA
from each stage using the Oligotex Direct mRNA kit
(Qiagen). mRNA was reverse transcribed into double
stranded cDNA using a modified oligo18(dT) primer with an
identifying tag sequence (see table below). cDNA's from
different stages were pooled in equal amounts before
adaptor ligation. Tag identification when sequencing from
5', end: Stage 1 (young fruitlet) insert 18(A)TCGGG; Stage
2 (young fruitlet 1cm dia) insert 18(A)TGCTG; Stage 3
(young fruitlet 12cm dia) insert 18(A)TCGGT; Stage 4
(maturing fruit I) insert 18(A)TCCGA; Stage 5 (maturing
fruit II) insert 18(A)TCGGA; Stage 6 (mature fruit) insert
18(A)TGGCT; Tag identification when sequencing from 3',
end: Stage 1 (young fruitlet) CACGA18(T) insert; Stage 2
(young fruitlet 1cm dia) CAGCA18(T) insert; Stage 3 (young
fruitlet 12cm dia) ACCGA18(T) insert; Stage 4 (maturing
fruit I) TCCGA18(T) insert; Stage 5 (maturing fruit II)
TCCGA18(T) insert; Stage 6 (mature fruit) ACGCA18(T)
insert. Double stranded cDNAs were size selected (more
than 450 bp), adaptor with EcoRI adaptors at both ends
and then digested with NotI. The cDNAs were then
directionally cloned into EcoRI-NotI digested pBS II SK(+) a
phagemid vector(Stratagene). Identification of adaptors
and tags in 5'-end sequenced clones:
<Vector>.. TAAAGTT<End Vector><Start
EcoRI adaptor>GATATCGAATTCATTGTGTGGG <End
EcoRI adaptor><Start Insert>.. AAAAAAAAAAAAAAA<End
Insert> <Start Tag>TGGCA<End Tag><Start
NotI site/Vector>GGGGCGCCACGGCGG... The total number of
white colony forming units (cfu) in the primary library
before amplification was 2.1x10⁶ cfu (colony forming
units). The background of empty clones was less than 1%.
Inserts ranged from 0.5kb to 4 kb, as determined by PCR.
Purified plasmid DNA from the primary library was
converted to single-stranded circles and used as a
template for PCR amplification using the T7 and T3 priming
sites flanking the cloned cDNA inserts. The purified PCR
products, representing the entire cloned cDNA population,
were used as a driver for normalization. Hybridization
between the single-stranded library and the PCR products
was carried out for 44 hours at 30C. Unhybridized
single-stranded DNA circles were separated from hybridized
DNA rendered partially double-stranded and electroporated
into DH10B cells to generate the normalized library. The
total number of clones with insert was 5.6x10⁶ cfu.
Background of empty clones was less than 1%.

FEATURES
source

```

ORIGIN
Query Match 74.5%; Score 16.4; DB 7; Length 555;
Best Local Similarity 94.4%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCACCTGAACGATAGCG 18
|||||

```

```
Db 86 CCACCTGAAGATAAGCG 69

RESULT 39
BM581827
LOCUS
DEFINITION
17000687274840 A.Gam.ad.cDNA blood1 Anopheles gambiae cDNA clone
19600449712152 5', mRNA sequence.
ACCESSION
BM581827
VERSION
BM581827.1 GI:18870294
KEYWORDS
EST.
SOURCE
Anopheles gambiae (African malaria mosquito)
ORGANISM
Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
Anopheles.
REFERENCE
1 (bases 1 to 644)
Holt R.A., Lin J.-J., Murphy S.D., Evans C.A., Kraft C.L.,
Charlab R., Collins F.H., Venter J.C. and Hoffman S.L.
Celera Anopheles gambiae EST project
Unpublished (2002)
CONTACT: Holt R.A.
Celera Genomics
45 W. Gude Dr., Rockville, MD 20850, USA
Tel: 2404533151
Fax: 2404534580
Email: HoltRA@celera.com
Plate: NU01004AWS row: C column: 18
Seq primer: M13 Reverse.
FEATURES
source
1..644
/organism="Anopheles gambiae"
/mol_type="mRNA"
/strain="RSP-ST (Reduced susc. to Permethrin - std.
chromosome)"
/db_xref="taxon:7165"
/clone="19600449712152"
/dev_stage="Adult"
/lab_host="DH10b"
/clone_lib="A.Gam.ad.cDNA.blood1"
/notes="Vector: pSport1; Site 1: SalI; Site 2: NotI; Whole
adult mosquitoes (mixed sex) - frozen on liquid nitrogen 24
hours after human blood feeding. cDNA inserts >500 bp
cloned directionally into pSport 1. Not 1 site is 3'.
Clones available through the Malaria Research and
Reference Reagent Resource Center (www.malaria.mr4.org)"
ORIGIN
Query Match 74.5%; Score 16.4; DB 4; Length 644;
Best Local Similarity 94.4%; Pred. No. 1.7e+03;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCACCTGAAGATAAGCG 18
||||| |||||||
Db 597 CCACCTAAGATAAGCG 614

RESULT 40
BX100177
LOCUS
DEFINITION
BX100177 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGp998J07469 ; IMAGE:232806, mRNA sequence.
ACCESSION
BX100177
VERSION
BX100177.1 GI:27830469
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 649)
Ebert L., Heil O., Hennig S., Neubert P., Patsch E., Peters M.,
Radelof U., Schneider D. and Korn B.
Human UniGeneSet - RZPD;
JOURNAL
COMMENT
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998J07469.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UniGeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/responderlibNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.
JOURNAL
COMMENT
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD; IMAGp998J07469.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UniGeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/responderlibNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.
FEATURES
Location/Qualifiers
1..649
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998J07469 ; IMAGE:232806"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAAGATTAAATTAAGATCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 04:22:34 ; Search time 166.757 Seconds
(without alignments)
780.981 Million cell updates/sec

Title: US-09-674-277-22

Perfect score: 22

Sequence: 1 cacccttccttcacccacac 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn19808:*
2: Geneseqn19908:*
3: Geneseqn20008:*
4: Geneseqn20018:*
5: Geneseqn20018b:*
6: Geneseqn20028a:*
7: Geneseqn20028b:*
8: Geneseqn20038a:*
9: Geneseqn20038b:*
10: Geneseqn20038c:*
11: Geneseqn20038d:*
12: Geneseqn20048a:*
13: Geneseqn20048b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed; and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	3	Aaz36122 Primer de
C 2	22	100.0	1181	3	Aaz36102 Nucleic a
C 3	20	90.9	341	8	Abx51595 Bovine ES
4	18.8	85.5	186391	11	Acn43938 Human gen
5	18.4	83.6	681	8	Abz52041 Aspergill
6	17.8	80.9	121	12	Adk91271 Polynucle
7	17.8	80.9	592	2	Aav88438 EST clone
8	17.8	80.9	1001	3	Aac57818 Arachidon
C 9	17.8	80.9	1206	12	Adj76014 Marker ge
10	17.8	80.9	4997	12	Adj40167 Plant cDN
11	17.8	80.9	12035	10	Adf65864 T-cell re
C 12	17.8	80.9	51698	12	Adg97614 Mouse can
13	17.8	80.9	68193	12	Adg97400 Mouse can
C 14	17.8	80.9	110000	12	Continuation (20 o
15	17.8	80.9	110000	12	Continuation (2 of
C 16	17.8	80.9	110000	12	Continuation (20 o
17	17.8	80.9	110000	12	Continuation (2 of
C 18	17.8	80.9	110000	12	Continuation (2 of
19	17.8	80.9	110000	12	Continuation (2 of
C 20	17.4	79.1	1530	6	Abz13748 Arabidops

C 21	17.4	79.1	1830	12	ADN13924	ADN13924 Human pro
C 22	17.4	79.1	1846	3	AAC34994	AAC34994 Arabidops
C 23	17.4	79.1	1920	12	ADQ67487	ADQ67487 Novel hum
C 24	17.4	79.1	32572	13	ADS36470	ADS36470 Human aut
C 25	17.4	79.1	86149	12	ADQ97278	ADQ97278 Human can
C 26	17.4	79.1	165221	11	ACN44524	ACN44524 Mouse gen
C 27	17.2	78.2	201	13	ADQ48103	ADQ48103 Myocardia
C 28	17.2	78.2	201	13	ADQ48232	ADQ48232 Myocardia
C 29	17.2	78.2	201	13	ADQ48173	ADQ48173 Myocardia
C 30	17.2	78.2	201	13	ADQ48254	ADQ48254 Myocardia
C 31	17.2	78.2	201	13	ADQ48102	ADQ48102 Myocardia
C 32	17.2	78.2	201	13	ADQ48136	ADQ48136 Myocardia
C 33	17.2	78.2	201	13	ADQ48141	ADQ48141 Myocardia
C 34	17.2	78.2	201	13	ADQ48331	ADQ48331 Myocardia
C 35	17.2	78.2	201	13	ADQ48231	ADQ48231 Myocardia
C 36	17.2	78.2	201	13	ADQ48274	ADQ48274 Myocardia
C 37	17.2	78.2	201	13	ADQ48129	ADQ48129 Myocardia
C 38	17.2	78.2	201	13	ADQ48250	ADQ48250 Myocardia
C 39	17.2	78.2	201	13	ADQ48333	ADQ48333 Myocardia
C 40	17.2	78.2	201	13	ADQ48193	ADQ48193 Myocardia
C 41	17.2	78.2	201	13	ADQ48194	ADQ48194 Myocardia
C 42	17.2	78.2	201	13	ADQ48296	ADQ48296 Myocardia
C 43	17.2	78.2	201	13	ADQ48294	ADQ48294 Myocardia
C 44	17.2	78.2	201	13	ADQ48312	ADQ48312 Myocardia
C 45	17.2	78.2	201	13	ADQ48314	ADQ48314 Myocardia

ALIGNMENTS

RESULT 1

Aaz36122

ID Aaz36122 standard; DNA; 22 BP.

XX Aaz36122;

XX 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virk gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

XX WO9955908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

PT Escherichia coli, particularly serotype O157:H7, used for detecting these

PT bacteria in food.

PS Claim 5; Page 27; 48pp; French.

XX Aaz36103-27 represent fragments derived from nucleic acid sequences
specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
derived from two sequences. The first (Aaz36101) is 99.9% homologous to
the katP gene of E. coli O157:H7 (nucleotides 407-1489 of Aaz36101), and
95.8% homologous with IS91 of E. coli (nucleotides 1-406 of Aaz36102).

XX The second sequence (Aaz36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 22 BP; 4 A; 11 C; 1 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 3; Length 22;
 Best Local Similarity 100.0%; Pred. No. 2.9;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCCAGAC 22
 |||||
 Db 1 CACCTTCCTTCATCTCCAGAC 22

RESULT 2

AAZ36102/c
 ID AAZ36102 standard; DNA; 1181 BP.

XX AC AAZ36102;

XX DT 11-FEB-2000 (first entry)

XX Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

XX enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.

XX Escherichia coli.

XX WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX DR WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.

XX PS Claim 1; Fig 2; 48pp; French.

XX The present sequence is specific to enterohemorrhagic Escherichia coli
 CC (EHEC). The sequence associated with the presence of virulence factors.
 CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
 CC 68% homology with the virK gene which codes for virulence proteins of
 CC Shigella flexneri. The present sequence is of plasmid origin. Fragments
 CC of the present sequence are used, as probes and primers, for detection of
 CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
 CC animal samples, foods or the environment. The fragments are also useful
 CC for epidemiological studies

XX SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 3; Length 1181;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCCAGAC 22
 |||||
 Db 1099 CACCTTCCTTCATCTCCAGAC 1078

RESULT 3

ABX51595/c

ID ABX51595 standard; cDNA; 341 BP.

XX AC ABX51595;

XX DT 25-FEB-2003 (first entry)

XX Bovine EST associated with lactation/muscle/fat deposition #1524.

XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

XX muscle deposition; fat deposition; genome mapping; gene identification;

XX gene analysis; cattle breeding.

XX Bos Taurus.

XX US2002137160-A1.

XX PD 26-SEP-2002.

XX PF 26-OCT-2001; 2001US-00983965.

XX PR 17-DEC-1998; 98US-0113678P.

XX PR 15-DEC-1999; 99US-00465231.

XX (BYAT/) BYATT J C.

XX (MATH/) MATHIALAGAN N.

XX (TAON/) TAO N.

XX (WARR/) WARREN W C.

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

XX WPI; 2003-102386/09.

XX Purified nucleic acid molecules, useful for genome mapping, gene
 PT identification and analysis, cattle breeding or preparation of constructs
 PT for cattle gene expression and genetically improved cattle.

XX PS Claim 2; SEQ ID NO 1524; 38pp; English.

XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 5912 nucleotide sequences,
 CC appearing as ABX50072-ABX55983, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non-translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 5912 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid; where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the 5912
 CC bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The present
 CC sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137160

XX SQ Sequence 341 BP; 83 A; 86 C; 97 G; 73 T; 0 U; 2 Other;

Query Match 90.9%; Score 20; DB 8; Length 341;
 Best Local Similarity 100.0%; Pred. No. 30;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCTCTTCCATCTCTCAG 22
 |||||
 Db 265 CCTCTTCCATCTCTCAG 246

RESULT 4

ACN43938
 ID ACN43938 standard; DNA; 186391 BP.

XX AC

XX ACN43938;

XX DT 18-NOV-2004 (first entry)

XX DE Human genomic sequence hCG40117.

XX KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

OS Homo sapiens.

XX WO2003073826-A2.

XX PD 12-SEP-2003.

XX PF 28-FEB-2003; 2003WO-US006235.

XX PR 01-MAR-2002; 2002US-00087192.

XX PA (SAGR-) SAGRES DISCOVERY.

XX PI Morris DW;

XX DR WPI; 2003-328604/31.

XX PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.

XX PS Claim 1; SEQ ID NO 136; Opp; English.

XX CC The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

XX SQ Sequence 186391 BP; 48359 A; 41525 C; 42262 G; 53708 T; 0 U; 537 Other;

Query Match 85.5%; Score 18.8; DB 11; Length 186391;
 Best Local Similarity 90.9%; Pred. No. 1.9e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCTCTTCCATCTCTCAG 22

Db 21054 CACCTTCTCTTCCATCTCTCAG 21075

RESULT 5

ABZ52041

ID ABZ52041 standard; cDNA; 681 BP.

XX AC

XX ABZ52041;

XX DT 28-MAR-2003 (first entry)

XX

DE

XX

KW

XX

OS

XX

PN

XX

PD

XX

PF

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Aspergillus oryzae polynucleotide SEQ ID NO 1154.

Aspergillus oryzae; fermentation; fungus; industrial; EST;

expressed sequence tag; gene; ss.

Aspergillus oryzae.

WO200279476-A1.

10-OCT-2002.

22-MAR-2002; 2002WO-1B000890.

30-MAR-2001; 2001JP-00098371.

(NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

(NARE-) NAT RES INST BREWING.

(NORQ) NAT FOOD RES INST MIN AGRIC.

Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;

Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;

WPI; 2003-046817/04.

Detection of expression of specific Aspergillus genes for monitoring the fermentation and growth conditions of the fungus, using DNA probes.

Claim 1; SEQ ID NO 1154; 48pp + Sequence Listing; Japanese.

The invention relates to a polynucleotide having any of 6006 specific sequences (ABZ50888-ABZ56893), which are expressed by a fungus under specific culture conditions including one or more of eutrophic, oligotrophic, solid, early germination, alkaline, high temperature, low temperature or maltose culture or polynucleotides stringently hybridizing to these sequences. The polynucleotides are useful for monitoring the progress of fermentation and the growth conditions of a fungus, especially of Aspergillus oryzae which is widely used in industrial fermentation. Also monitoring for fungal contamination. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 681 BP; 149 A; 210 C; 185 G; 137 T; 0 U; 0 Other;

Query Match 83.6%; Score 18.4; DB 8; Length 681;

Best Local Similarity 95.0%; Pred. No. 1.7e+02;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CACCTTCTCTTCCATCTCTCAG 20

Db 417 CACCTTCTCTTCCATCTCTCAG 436

RESULT 6

ADK91271

ID ADK91271 standard; DNA; 121 BP.

XX ADK91271;

XX 06-MAY-2004 (first entry)

Polynucleotide used to detect SNPs of the invention #300.

human; single nucleotide polymorphism; SNP; ds.

OS Homo sapiens.

XX JP2003259875-A.

XX 16-SEP-2003.

XX 08-MAR-2002; 2002JP-00064373.

XX 08-MAR-2002; 2002JP-00064373.
 XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 XX WPI; 2004-093977/10.
 DR Novel polynucleotide useful for PCR amplification along with two DNA
 PT fragment from another set of sequences, or for detecting single
 PT nucleotide polymorphism in human gene.
 XX Claim 1; SEQ ID NO 300; 2627pp; Japanese.
 XX The present invention relates to a polynucleotide isolated from a human
 CC gene and is useful for detecting a single nucleotide polymorphism in a
 CC human gene or for diagnosing of disease. The invention enables the
 CC detection of a single nucleotide polymorphism in a human gene. The
 CC present sequence represents a polynucleotide used to detect SNPs of the
 CC invention.
 XX Sequence 121 BP; 22 A; 45 C; 30 G; 23 T; 0 U; 1 Other;
 SQ

Query Match 80.9%; Score 17.8; DB 12; Length 121;
 Best Local Similarity 90.5%; Pred. No. 2.6e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CACCTTCCTTCCTCCATCCTCAGA 21
 ||| ||||| ||||| |||||
 DB 20 CACCCTCCTTCCTCCATCCTCAGA 40

RESULT 7
 AAV88438
 ID AAV88438 standard; cDNA; 592 BP.
 XX AAV88438;
 AC AAV88438;
 XX 12-FEB-1999 (first entry)
 DT EST clone GP218.
 XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 XX Homo sapiens.
 OS WO9845437-A2.
 FN 15-OCT-1998.
 PD 10-APR-1998; 98WO-US006956.
 PF 10-APR-1997; 97US-00837312.
 PR (GEMY) GENETICS INST INC.
 XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 PI Spaulding V, Agostino M;
 XX WPI; 1999-070078/06.
 DR New polynucleotides encoding human secreted proteins - derived from e.g.
 XX human blood, kidney, foetal lung, placenta, testes, brain, ovary,
 PT pituitary, retina and colon cDNA libraries.
 XX Claim 1; Page 397; 641pp; English.
 PS The present sequence represents an expressed sequence tag (EST), and is a
 CC polynucleotide of the invention. The polynucleotides of the invention are
 CC all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to

CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene therapy
 XX Sequence 592 BP; 172 A; 156 C; 135 G; 128 T; 0 U; 1 Other;
 SQ

Query Match 80.9%; Score 17.8; DB 2; Length 592;
 Best Local Similarity 90.5%; Pred. No. 3e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CACCTTCCTTCCTCCATCCTCAGA 21
 ||| ||||| ||||| |||||
 DB 108 CAGCTTCCTTCCTCCATCCTCAGA 128

RESULT 8
 AAC57818
 ID AAC57818 standard; DNA; 1001 BP.
 XX AAC57818;
 AC AAC57818;
 XX 25-JAN-2001 (first entry)
 DT Arachidonic acid metabolism related genomic biallelic marker #452.
 DE Human; biallelic marker; arachidonic acid metabolism; genotyping;
 KW detection; hybridisation; phenotype; haplotype; SNP; polymorphic base;
 KW single nucleotide polymorphism; hybridisation assay; sequencing assay;
 KW specific amplification assay; identification; ERBM; 12-LO-RBM;
 KW eicosanoid-related biallelic marker; 12-LO-related biallelic marker; ds.
 XX Homo sapiens.
 OS WO200047771-A2.
 FN 17-AUG-2000.
 PD 11-FEB-2000; 2000WO-IB000184.
 PF 12-FEB-1999; 99US-0119917P.
 PR 23-MAR-1999; 99US-00275267.
 PR 07-MAY-1999; 99US-0133200P.
 XX (GEST) GENSET.
 PA Blumenfeld M, Bougueleret L, Chumakov I;
 PI WPI; 2000-571881/53.
 DR Novel biallelic markers useful for detecting conditions and genotypes
 XX associated with arachidonic acid metabolism.
 PT Claim 13; Page 613; 802pp; English.
 PS The present invention describes polynucleotides including biallelic
 CC markers derived from genes involved in arachidonic acid metabolism and
 CC from genomic regions flanking those genes. Methods from the present
 CC invention may be used to select individuals for clinical trials and
 CC predict responses to treatment with drugs. The polynucleotides may be
 CC used in hybridisation assays, sequencing assays and specific
 CC amplification assays for identifying an eicosanoid-related biallelic
 CC marker (ERBM) or 12-LO-related biallelic marker, and for amplifying a
 CC segment of nucleotides containing an ERBM. The polynucleotides are useful
 CC in diagnostic kits. The markers may be used to detect conditions and
 CC genotypes associated with arachidonic acid metabolism. AAC57367 to
 CC AAC58018 and AAB24019 and AAB24020 represent sequences used in the
 CC exemplification of the present invention. N.B. Polymorphic bases (single

CC nucleotide polymorphisms also known as SNPs) in the polynucleotide
CC sequences from the present invention have been given as their
CC corresponding degenerate bases e.g. a polymorphic base of C or T has been
CC given as Y
XX

SQ Sequence 1001 BP; 305 A; 215 C; 190 G; 290 T; 0 U; 1 Other;

Query Match 80.9%; Score 17.8; DB 3; Length 1001;

Best Local Similarity 90.5%; Pred. No. 3.2e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCCTCAGA 21

|||||

Db 759 CATCTTCCTTCCTTCCTCCTCAGA 779

RESULT 9

ADJ76014/c

ID ADJ76014 standard; DNA; 1206 BP.

XX AC ADJ76014;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene SEQ ID NO:1266.

XX KW bronchial asthma; chronic obstructive pulmonary disease;

XX KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

XX KW gene therapy; marker gene; gene; ds.

XX OS Mus musculus.

XX PN EP1394274-A2.

XX PD 03-MAR-2004.

XX PF 04-AUG-2003; 2003EP-00254857.

XX PR 06-AUG-2002; 2002JP-00229312.

XX PR 20-MAR-2003; 2003JP-00077212.

XX PA (GENO-) GENOX RES INC.

XX PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;

XX DR WPI; 2004-193155/19.

XX PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX PT comparing the expression level of a marker gene in a biological sample
XX PT from a subject with the expression level of the gene in a sample from a
XX PT healthy subject.

XX PS Example 11; SEQ ID NO 1266; 241pp; English.

XX CC The present invention describes a method of testing for bronchial asthma
XX CC or chronic obstructive pulmonary disease. The method comprises
XX CC determining the expression level of a marker gene in a biological sample
XX CC from a subject, comparing the expression level determined with the
XX CC expression level of the marker gene in a biological sample from a healthy
XX CC subject, and judging whether the subject has bronchial asthma or chronic
XX CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
XX CC genes (S1) whose expression levels increase when respiratory epithelial
XX CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
XX CC whose expression levels decrease when respiratory epithelial cells are
XX CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
XX CC testing for bronchial asthma or chronic obstructive pulmonary disease;
XX CC (2) a kit for screening for a candidate compound for a therapeutic agent
XX CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
XX CC an animal model for bronchial asthma or chronic obstructive pulmonary
XX CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
XX CC method for producing an animal model for bronchial asthma or chronic
XX CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
XX CC asthma or chronic obstructive pulmonary disease, comprising the compound,

CC a marker gene or an antisense nucleic acid corresponding to a portion of
CC the marker gene, a ribozyme, a polynucleotide that suppresses the
CC expression of the gene through an RNAi effect or an antibody recognising
CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC probe has been immobilised to assay a marker gene. (I) has respiratory
CC and antiasthmatic activities, and can be used in gene therapy. The method
CC is useful for testing for or screening for a therapeutic agent for
CC bronchial asthma or chronic obstructive pulmonary disease. The present
CC sequence is used in the exemplification of the present invention.

XX SQ Sequence 1206 BP; 345 A; 271 C; 284 G; 306 T; 0 U; 0 Other;

Query Match 80.9%; Score 17.8; DB 12; Length 1206;

Best Local Similarity 90.5%; Pred. No. 3.3e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCCTCAGA 21

|||||

Db 450 CTCCTTCCTTCCTCCTCCTCAGA 430

RESULT 10

ADJ40167

ID ADJ40167 standard; cDNA; 4997 BP.

XX AC ADJ40167;

XX DT 06-MAY-2004 (first entry)

XX DE Plant cDNA #1167.

XX KW Plant; gene; ss; transcription; plant genome augmentation; cereal;

XX KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;

XX KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;

XX KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;

XX KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;

XX KW antifungal.

XX OS Eukaryota.

XX PN US2004016025-A1.

XX PD 22-JAN-2004.

XX PF 26-SEP-2002; 2002US-00260238.

XX PR 26-SEP-2001; 2001US-0325277P.

XX PR 26-SEP-2001; 2001US-0325448P.

XX PR 04-APR-2002; 2002US-0370620P.

XX PA (BUDW/) BUDWORTH P.

XX PA (MOUG/) MOUGHAMER T.

XX PA (BRIG/) BRIGGS S P.

XX PA (COOP/) COOPER B.

XX PA (GLAZ/) GLAZERBROOK J.

XX PA (GOFF/) GOFF S A.

XX PA (KATA/) KATAGIRI F.

XX PA (KREP/) KREPS J.

XX PA (PROV/) PROVANT N.

XX PA (RICK/) RICKE D.

XX PA (ZHUT/) ZHU T.

XX PI Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;

XX PI Goff SA, Katagiri F, Kreps J, Provant N, Ricke D, Zhu T;

XX DR WPI; 2004-190374/18.

XX PT New rice promoter, useful for manipulating crop plants to alter or
XX PT improve phenotypic characteristics, e.g. produce large quantities of oil
XX PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
XX PT or high nutritional value.

PS Claim 25; SEQ ID NO 1167; 230pp; English.

XX The invention relates to plant nucleotide sequences that direct seed-,

CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential

CC or constitutive transcription of an operatively linked nucleic acid

CC segment. The invention also relates to a method for augmenting a plant

CC genome and a method of identifying a gene, where its expression is

CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive

CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,

CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,

CC sorghum, rice or wheat. The polynucleotides and the polypeptides they

CC encode are useful for manipulating crop plants to alter or improve

CC phenotypic characteristics, to produce large quantities of oil or

CC proteins, to incur resistance to insecticides, viruses or fungi, and to

CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants

CC have a high nutritional value with reduced apical dominance or dwarfism,

CC early flowering or altered metabolic pathways. This sequence represents a

CC plant nucleic acid of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification but was obtained in

CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 4997 BP; 1373 A; 1057 C; 1246 G; 1316 T; 0 U; 5 Other;

Query Match 80.9%; Score 17.8; DB 12; Length 4997;

Best Local Similarity 90.5%; Pred. No. 3.8e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCCTCAGA 21

4666 CACCTTCCTTCCTCCTCCTCAGA 4686

Db

RESULT 11

ADF65864

ID ADF65864 standard; DNA; 12035 BP.

XX

AC ADF65864;

XX

XX 12-FEB-2004 (first entry)

DT

DE T-cell receptor protein related DNA, SEQ ID 8.

XX

XX Immunoregulator; natural killer T-cell; NKT; T-cell receptor;

XW Th2 type cytokine; murine; ds.

XX

OS Mus musculus.

XX

XX JP2003199587-A.

PN

XX

PD 15-JUL-2003.

XX

XX 27-SEP-2002; 2002JP-00283379.

PF

XX

XX 27-SEP-2001; 2001JP-00297452.

PR

XX (MITU) MITSUBISHI CHEM CORP.

PA

XX WPI; 2003-819598/77.

DR

XX

XX Novel V-alpha19 natural killer T-cell comprises T-cell receptor, produces

PT Th2 type cytokine in response to stimulation through T-cell receptor.

PT

XX

PS Claim 14; SEQ ID NO 8; 27pp; Japanese.

XX

XX The present invention relates to natural killer T-cells (NKT; I)

CC comprising T-cell receptor (T1), which has a 128 amino acid sequence (S1;

CC ADF65857). (I) produces Th2 type cytokine in response to a stimulation

CC through the T-cell receptor. Also claimed is an antibody (II) which

CC recognises (I). (II) is useful as active ingredient of immunoregulatory

CC agent and for acquiring NKT cell and for measuring NKT cell in blood.

CC (II) is also useful to screen a test substance, which has

CC immunoregulatory activity.

XX

SQ Sequence 12035 BP; 3245 A; 2981 C; 2803 G; 3006 T; 0 U; 0 Other;

Query Match 80.9%; Score 17.8; DB 10; Length 12035;

Best Local Similarity 90.5%; Pred. No. 4.1e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCCTCAGA 21

330 CACCTTCCTTCCTCCTCCTCAGA 350

Db

RESULT 12

ADQ97614/c

ID ADQ97614 standard; DNA; 51698 BP.

XX

AC ADQ97614;

XX

DT 07-OCT-2004 (first entry)

DT

XX

DE Mouse cancer associated sequence MD10-015, SEQ ID 591.

XX

XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

XX

OS Mus musculus.

XX

XX WO2004060304-A2.

PN

XX 22-JUL-2004.

PD

XX 22-DEC-2003; 2003WO-US041389.

PF

XX 27-DEC-2002; 2002US-00330773.

PR

XX (SAGR-) SAGRES DISCOVERY INC.

PA

XX Morris DW, Malandro MS;

PI

XX WPI; 2004-543781/52.

DR

XX

XX New isolated cancer associated nucleic acids comprising at least 10

PT contiguous nucleotides, useful for diagnosing, preventing and/or treating

PT cancers such as leukemia and lymphoma.

XX

XX Claim 1; SEQ ID NO 591; 199pp; English.

PS

XX The present invention relates to cancer associated sequences (ADQ97025-

CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or

CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence

CC data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 51698 BP; 13873 A; 11031 C; 12067 G; 14727 T; 0 U; 0 Other;

Query Match 80.9%; Score 17.8; DB 12; Length 51698;

Best Local Similarity 90.5%; Pred. No. 4.8e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCCTCAGA 21

19065 CACCTTCCTTCCTCCTCCTCAGA 19045

Db

RESULT 13

ADQ97400

ID ADQ97400 standard; DNA; 68193 BP.

XX

AC ADQ97400;

XX

DT 07-OCT-2004 (first entry)

DT

XX

DE Mouse cancer associated sequence MD08-044, SEQ ID 377.

XX

KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.

XX Mus musculus.

XX WO2004060304-A2.

XX 22-JUL-2004.

XX 22-DEC-2003; 2003WO-US041389.

XX 27-DEC-2002; 2002US-00330773.

XX (SAGR-) SAGRES DISCOVERY INC.

XX Morris DW, Malandro MS;

XX WPI; 2004-543781/52.

XX New isolated cancer associated nucleic acids comprising at least 10 contiguous nucleotides, useful for diagnosing, preventing and/or treating cancers such as leukemia and lymphoma.

XX Claim 1; SEQ ID NO 377; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-ADQ98004). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 68193 BP; 18706 A; 14414 C; 14071 G; 20571 T; 0 U; 431 Other;

Query Match 80.9%; Score 17.8; DB 12; Length 68193;

Best Local Similarity 90.5%; Pred. No. 4.9e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCAGA 21

DB 58078 CTCCTTCCTTCCTCCTCAGA 58098

RESULT 14

ADN46845_19/c

Continuation (20 of 21) of ADN46845 from base 1900001 (Thermococcus kodakaraensis KOD1 Sequence split into 21 fragments LOCUS ADN46845 Accession Adn46845

WP Fragment Name Begin End

WP ADN46845_00 1 110000

WP ADN46845_01 100001 210000

WP ADN46845_02 200001 310000

WP ADN46845_03 300001 410000

WP ADN46845_04 400001 510000

WP ADN46845_05 500001 610000

WP ADN46845_06 600001 710000

WP ADN46845_07 700001 810000

WP ADN46845_08 800001 910000

WP ADN46845_09 900001 1010000

WP ADN46845_10 1000001 1110000

WP ADN46845_11 1100001 1210000

WP ADN46845_12 1200001 1310000

WP ADN46845_13 1300001 1410000

WP ADN46845_14 1400001 1510000

WP ADN46845_15 1500001 1610000

WP ADN46845_16 1600001 1710000

WP ADN46845_17 1700001 1810000

WP ADN46845_18 1800001 1910000

WP ADN46845_19 1900001 2010000

WP ADN46845_20 2000001 2089378

Query Match 80.9%; Score 17.8; DB 12; Length 110000;

Best Local Similarity 90.5%; Pred. No. 5.2e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCAGA 21

DB 30919 CACCTTCCTTCCTCCTCAGAA 30899

RESULT 15

ADN47591_01

Continuation (2 of 21) of ADN47591 from base 100001 (Thermococcus kodakaraensis KOD1 Sequence split into 21 fragments LOCUS ADN47591 Accession Adn47591

WP Fragment Name Begin End

WP ADN47591_00 1 110000

WP ADN47591_01 100001 210000

WP ADN47591_02 200001 310000

WP ADN47591_03 300001 410000

WP ADN47591_04 400001 510000

WP ADN47591_05 500001 610000

WP ADN47591_06 600001 710000

WP ADN47591_07 700001 810000

WP ADN47591_08 800001 910000

WP ADN47591_09 900001 1010000

WP ADN47591_10 1000001 1110000

WP ADN47591_11 1100001 1210000

WP ADN47591_12 1200001 1310000

WP ADN47591_13 1300001 1410000

WP ADN47591_14 1400001 1510000

WP ADN47591_15 1500001 1610000

WP ADN47591_16 1600001 1710000

WP ADN47591_17 1700001 1810000

WP ADN47591_18 1800001 1910000

WP ADN47591_19 1900001 2010000

WP ADN47591_20 2000001 2089378

Query Match 80.9%; Score 17.8; DB 12; Length 110000;

Best Local Similarity 90.5%; Pred. No. 5.2e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCAGA 21

DB 58459 CACCTTCCTTCCTCCTCAGAA 58479

RESULT 16

ADN46123_19/c

Continuation (20 of 21) of ADN46123 from base 1900001 (Thermococcus kodakaraensis KOD1 Sequence split into 21 fragments LOCUS ADN46123 Accession Adn46123

WP Fragment Name Begin End

WP ADN46123_00 1 110000

WP ADN46123_01 100001 210000

WP ADN46123_02 200001 310000

WP ADN46123_03 300001 410000

WP ADN46123_04 400001 510000

WP ADN46123_05 500001 610000

WP ADN46123_06 600001 710000

WP ADN46123_07 700001 810000

WP ADN46123_08 800001 910000

WP ADN46123_09 900001 1010000

WP ADN46123_10 1000001 1110000

WP ADN46123_11 1100001 1210000

WP ADN46123_12 1200001 1310000

WP ADN46123_13 1300001 1410000

WP ADN46123_14 1400001 1510000

WP ADN46123_15 1500001 1610000

WP ADN46123_16 1600001 1710000

WP ADN46123_17 1700001 1810000

WP ADN46123_18 1800001 1910000

WP ADN46123_19 1900001 2010000

WP ADN46123_20 2000001 2089378

Query Match 80.9%; Score 17.8; DB 12; Length 110000;

Best Local Similarity 90.5%; Pred. No. 5.2e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCAGA 21

Db 30919 CACCTTCCTTCCTCCTCTCAAA 30899

RESULT 17

ADN47209_01

Continuation (2 of 21) of ADN47209 from base 100001 (Thermococcus kodakaraensis KOD1 gen

WP Sequence split into 21 fragments LOCUS ADN47209 Accession Adn47209

WP Fragment Name Begin End

WP ADN47209_00 1 110000

WP ADN47209_01 100001 210000

WP ADN47209_02 200001 310000

WP ADN47209_03 300001 410000

WP ADN47209_04 400001 510000

WP ADN47209_05 500001 610000

WP ADN47209_06 600001 710000

WP ADN47209_07 700001 810000

WP ADN47209_08 800001 910000

WP ADN47209_09 900001 1010000

WP ADN47209_10 1000001 1110000

WP ADN47209_11 1100001 1210000

WP ADN47209_12 1200001 1310000

WP ADN47209_13 1300001 1410000

WP ADN47209_14 1400001 1510000

WP ADN47209_15 1500001 1610000

WP ADN47209_16 1600001 1710000

WP ADN47209_17 1700001 1810000

WP ADN47209_18 1800001 1910000

WP ADN47209_19 1900001 2010000

WP ADN47209_20 2000001 2089378

Query Match 80.9%; Score 17.8; DB 12; Length 110000;

Best Local Similarity 90.5%; Pred. No. 5.2e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCTCAAA 58479

Db 58459 CACCTTCCTTCCTCCTCTCAAA 58479

RESULT 18

ADN46464_19/c

Continuation (20 of 21) of ADN46464 from base 1900001 (Thermococcus kodakaraensis KOD1 g

WP Sequence split into 21 fragments LOCUS ADN46464 Accession Adn46464

WP Fragment Name Begin End

WP ADN46464_00 1 110000

WP ADN46464_01 100001 210000

WP ADN46464_02 200001 310000

WP ADN46464_03 300001 410000

WP ADN46464_04 400001 510000

WP ADN46464_05 500001 610000

WP ADN46464_06 600001 710000

WP ADN46464_07 700001 810000

WP ADN46464_08 800001 910000

WP ADN46464_09 900001 1010000

WP ADN46464_10 1000001 1110000

WP ADN46464_11 1100001 1210000

WP ADN46464_12 1200001 1310000

WP ADN46464_13 1300001 1410000

WP ADN46464_14 1400001 1510000

WP ADN46464_15 1500001 1610000

WP ADN46464_16 1600001 1710000

WP ADN46464_17 1700001 1810000

WP ADN46464_18 1800001 1910000

WP ADN46464_19 1900001 2010000

WP ADN46464_20 2000001 2089378

Query Match 80.9%; Score 17.8; DB 12; Length 110000;

Best Local Similarity 90.5%; Pred. No. 5.2e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCTCAAA 58479

Db 30919 CACCTTCCTTCCTCCTCTCAAA 30899

RESULT 19

ADN47960_01

Continuation (2 of 21) of ADN47960 from base 100001 (Thermococcus kodakaraensis KOD1 gen

WP Sequence split into 21 fragments LOCUS ADN47960 Accession Adn47960

WP Fragment Name Begin End

WP ADN47960_00 1 110000

WP ADN47960_01 100001 210000

WP ADN47960_02 200001 310000

WP ADN47960_03 300001 410000

WP ADN47960_04 400001 510000

WP ADN47960_05 500001 610000

WP ADN47960_06 600001 710000

WP ADN47960_07 700001 810000

WP ADN47960_08 800001 910000

WP ADN47960_09 900001 1010000

WP ADN47960_10 1000001 1110000

WP ADN47960_11 1100001 1210000

WP ADN47960_12 1200001 1310000

WP ADN47960_13 1300001 1410000

WP ADN47960_14 1400001 1510000

WP ADN47960_15 1500001 1610000

WP ADN47960_16 1600001 1710000

WP ADN47960_17 1700001 1810000

WP ADN47960_18 1800001 1910000

WP ADN47960_19 1900001 2010000

WP ADN47960_20 2000001 2089378

Query Match 80.9%; Score 17.8; DB 12; Length 110000;

Best Local Similarity 90.5%; Pred. No. 5.2e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCTCAAA 58479

Db 58459 CACCTTCCTTCCTCCTCTCAAA 58479

RESULT 20

ABZ13748/c

ID ABZ13748 standard; DNA; 1530 BP.

XX ABZ13748;

XX 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 1553.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

XX 26-JAN-2001; 2001US-0264647P.

XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI) SCRIPPS RES INST.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

XX Identifying a stress condition to which a plant cell has been exposed and

XX producing plants with increased tolerance to these abiotic stresses.

PS Claim 144; SEQ ID NO 1553; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant cell has been exposed, comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office

XX

SQ Sequence 1530 BP; 403 A; 316 C; 390 G; 421 T; 0 U; 0 Other;

Query Match 79.1%; Score 17.4; DB 6; Length 1530;

Best Local Similarity 94.7%; Pred. No. 5.1e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTTCTTCATCCTCAGA 21
|||||

Db 711 CCTTCTGCATCCTCAGA 693
|||||

RESULT 21

ADN13924/c

ID ADN13924 standard; cDNA; 1830 BP.

XX

AC ADN13924;

XX

XX 29-JUL-2004 (first entry)

XX Human prostate/colon/lung/breast cancer-related cDNA 1439, SEQ:1439.

XX Human; cancer; tumour; prostate cancer; colon cancer; lung cancer; breast cancer; drug screening; diagnosis; prognosis; prevention; gene mapping; tissue typing; tissue profiling; cytostatic; gene therapy; Gene; ss.

XX Homo sapiens.

XX

XX Key Location/Qualifiers

FT CDS 1..1830

FT /tag= a

FT /product= "Human prostate/colon/lung/breast cancer-related protein 1439"

XX

XX WO2004039943-A2.

XX

XX 13-MAY-2004.

XX

XX 16-MAY-2003; 2003WO-US015465.

XX

XX 17-MAY-2002; 2002US-0381533P.

XX 04-FEB-2003; 2003US-0445222P.

XX (CHIR) CHIRON CORP.

XX

XX Scott EM, Lamson G, Kassam A, Zhang G, Sakamoto D, Garcia PD; WPI; 2004-376173/35.

XX P-PSDB; ADN13981.

XX

XX New isolated polynucleotides, useful for gene mapping or tissue typing or profiling, as diagnostic reagents, and for preventing or treating cancer, e.g. prostate, colon, or breast cancer.

XX

XX Claim 2; SEQ ID NO 1439; 190pp; English.

XX

XX The invention relates to nucleic acids (ADN12486-ADN13970) isolated from human prostate, colon, lung and breast cancer cDNA libraries, and to 57

CC proteins (ADN13971-ADN14027) encoded by a subset of these cDNA sequences (ADN13914-ADN13970). The invention also relates to vectors and host cells comprising a nucleic acid of the invention; a method for the recombinant production of a protein of the invention; an antibody specific for a protein of the invention; a polynucleotide library comprising at least one nucleic acid sequence of the invention; a method for detecting a cancerous cell by PCR or probe hybridisation; inhibiting a cancerous phenotype (particularly aberrant proliferation) of a cell; a method of identifying an agent that modulates the biological activity of a gene product differentially expressed in a cancerous cell compared with a normal cell; and a method of treating a cancer patient using the agent identified. The nucleic acids and polypeptides can be used to diagnose, prognose, treat or prevent cancers such as prostate, colon, lung or breast cancer, and can also be used to screen for drugs for the treatment of cancer. The nucleic acids can also be used for gene mapping, tissue typing and tissue profiling. The present sequence represents a specifically claimed cancer-related cDNA of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX

SQ Sequence 1830 BP; 484 A; 462 C; 522 G; 362 T; 0 U; 0 Other;

Query Match 79.1%; Score 17.4; DB 12; Length 1830;

Best Local Similarity 94.7%; Pred. No. 5.1e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTTCTTCATCCTCAGA 21
|||||

Db 882 CCTTCTTCACCTCAGA 864
|||||

RESULT 22

AAC34994/c

ID AAC34994 standard; DNA; 1846 BP.

XX

AC AAC34994;

XX

XX 17-OCT-2000 (first entry)

XX

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 8631.

XX

XX Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX

XX 06-SEP-2000.

XX

XX 25-FEB-2000; 2000EP-00301439.

XX

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 23-APR-1999; 99US-0130891P.

XX 28-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 30-APR-1999; 99US-0132407P.

XX 04-MAY-1999; 99US-0132484P.

XX 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134811P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135529P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140891P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142820P.
PR 12-JUL-1999; 99US-0142877P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.

PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-015753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 79.1%; Score 17.4; DB 3; Length 1846;
Best Local Similarity 94.7%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CTTCTTCCATCTCAGA 21
|||||
Db 769 CTTCTTCCATCTCAGA 751

RESULT 23
ADQ67487/c
ID ADQ67487 standard; cDNA; 1920 BP.

XX AC ADQ67487;
XX DE
DT 07-OCT-2004 (first entry)
XX DE
XX DE Novel human cDNA sequence #2460.
XX ss; gene; osteoprotective; neuroprotective; nontropic; antiparkinsonian;
KW cytotatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX Homo sapiens.

XX EP1440981-A2.
XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

XX 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Negai K, Irie R;

XX WPI; 2004-535376/52.

XX P-PSDB; ADQ67794.

XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

XX Claim 1; SEQ ID NO 4648; 2449pp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.

XX SQ Sequence 1920 BP; 530 A; 492 C; 415 G; 483 T; 0 U; 0 Other;

Query Match 79.1%; Score 17.4; DB 12; Length 1920;
Best Local Similarity 94.7%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTTCTTCCATCTCAGAC 22

Db 1311 CTTCTTCCATCTCAGAC 1293

RESULT 24
ADS36470/c
ID ADS36470 standard; DNA; 32572 BP.

XX AC ADS36470;

XX DT 16-DEC-2004 (first entry)

XX DE Human autoimmune disease-related genomic DNA sequence - SEQ ID 1684.

XX single nucleotide polymorphism detection; SNP detection;
KW rheumatoid arthritis; type 1 diabetes; multiple sclerosis;
KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;
KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;
KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;
KW primary systemic vasculitis; ds.

XX Homo sapiens.

XX WO2004083403-A2.

XX 30-SEP-2004.

XX 18-MAR-2004; 2004WO-US008461.

XX 18-MAR-2003; 2003US-0455444P.

XX 25-APR-2003; 2003US-0465241P.

XX (APPL-) APPLERA CORP.

XX Cargill M, Begovich AB, Alexander HC;

XX WPI; 2004-728480/71.

XX New isolated nucleic acid molecule comprises at least 8 contiguous
PT nucleotides where one of the nucleotides is a single nucleotide
PT polymorphism (SNP), useful for diagnosing or treating autoimmune
PT diseases, e.g. rheumatoid arthritis.

XX Claim 16; SEQ ID NO 1684; 123pp; English.

XX The invention comprises amino acid and coding sequences containing
CC genetic polymorphisms associated with an altered risk of developing an
CC autoimmune disease (e.g. rheumatoid arthritis). The invention further
CC comprises a method of identifying an individual that has an altered risk
CC of developing an autoimmune disease, comprising detecting a single
CC nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA
CC and protein sequences of the invention are useful for diagnosing and
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,
CC myocardiitis, Sjogren's disease, or primary systemic vasculitis. The
CC present nucleic acid represents a human autoimmune disease-related
CC genomic DNA sequence of the invention. NOTE: The present sequence is not
CC shown in the specification, but has been retrieved from the WIPO website.

XX SQ Sequence 32572 BP; 8336 A; 7467 C; 7930 G; 8760 T; 0 U; 79 Other;

Query Match 79.1%; Score 17.4; DB 13; Length 32572;
Best Local Similarity 94.7%; Pred. No. 6.9e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTTCTTCCATCCTCAGAC 22
|||||
Db 28857 CTTCTTCCATCCTCTGAC 28839
|||||

RESULT 25
ADQ97278/c
ID ADQ97278 standard; DNA; 86149 BP.
XX
AC ADQ97278;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human cancer associated sequence HD08-025, SEQ ID 254.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS Homo sapiens.
XX
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
SQ New isolated cancer associated nucleic acids comprising at least 10 contiguous nucleotides, useful for diagnosing, preventing and/or treating cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 254; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-ADQ98004). The sequences are useful for the diagnosis, prevention and/or treatment of cancer, such as leukemia and lymphoma. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 86149 BP; 19445 A; 19457 C; 18668 G; 20667 T; 0 U; 7912 Other;

Query Match 79.1%; Score 17.4; DB 12; Length 86149;
Best Local Similarity 94.7%; Pred. No. 7.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CTTCTTCCATCCTCAGA 21
|||||
Db 56733 CTTCTTCCATCCTCAGA 56715
|||||

RESULT 26
ACN44524/c
ID ACN44524 standard; DNA; 165221 BP.
XX
AC ACN44524;
XX
DT 18-NOV-2004 (first entry)
XX
DE Mouse genomic sequence mCG20599.
XX
KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
OS Mus musculus.

PN WO2003073826-A2.
XX
PD 12-SEP-2003.
XX
PF 28-FEB-2003; 2003WO-US006235.
XX
PR 01-MAR-2002; 2002US-00087192.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-328604/31.
XX
PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma comprises a nucleotide sequence.
XX
PS Claim 1; SEQ ID NO 1015; Opp; English.
XX
CC The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published
XX
SQ Sequence 165221 BP; 41378 A; 37617 C; 38634 G; 43238 T; 0 U; 4354 Other;

Query Match 79.1%; Score 17.4; DB 11; Length 165221;
Best Local Similarity 94.7%; Pred. No. 8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTTCTTCCATCCTCAGAC 22
|||||
Db 48021 CTTCTTCCATCCTCTGAC 48003
|||||

RESULT 27
ADQ48103/c
ID ADQ48103 standard; DNA; 201 BP.
XX
AC ADQ48103;
XX
DT 18-NOV-2004 (first entry)
XX
DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 9766.
XX
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
PD 15-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US040978.
XX
PR 20-DEC-2002; 2002US-0434778P.
XX
PR 10-MAR-2003; 2003US-0453135P.
XX
PR 30-APR-2003; 2003US-0466412P.
XX
PR 23-SEP-2003; 2003US-0504955P.
XX
PA (APPL-) APPLERA CORP.
XX

PI Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
XX
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
XX Claim 7; SEQ ID NO 9766; 145pp; English.
XX
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
XX downloaded from the WIPO website.
XX
SQ Sequence 201 BP; 55 A; 52 C; 62 G; 31 T; 0 U; 1 Other;

Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCTCAGAC 22
Db 158 CACCTTCCTTCCTCTCTCAGAC 137

RESULT 28
ADQ48232/c
ID ADQ48232 standard; DNA; 201 BP.
XX
XX ADQ48232;
XX
XX 18-NOV-2004 (first entry)
XX
XX Myocardial infarction-associated SNP flanking transcript, SEQ ID 9895.
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
XX Homo sapiens.
XX
XX WO2004058052-A2.
XX
XX 15-JUL-2004.
XX
XX 22-DEC-2003; 2003WO-US040978.
XX
XX 20-DEC-2002; 2002US-0434778P.
XX
XX 10-MAR-2003; 2003US-0453135P.

PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX
XX (APPL-) APPLERA CORP.
XX
XX Cargill M, Devlin JJ, Iakoubova O;
XX WPI; 2004-533949/51.
XX
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
XX Claim 7; SEQ ID NO 9895; 145pp; English.
XX
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
XX downloaded from the WIPO website.
XX
SQ Sequence 201 BP; 55 A; 52 C; 62 G; 31 T; 0 U; 1 Other;

Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCTCAGAC 22
Db 158 CACCTTCCTTCCTCTCTCAGAC 137

RESULT 29
ADQ48173/c
ID ADQ48173 standard; DNA; 201 BP.
XX
XX ADQ48173;
XX
XX 18-NOV-2004 (first entry)
XX
XX Myocardial infarction-associated SNP flanking transcript, SEQ ID 9836.
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
XX Homo sapiens.
XX
XX WO2004058052-A2.
XX
XX 15-JUL-2004.

XX 22-DEC-2003; 2003WO-US040978.
 XX PF
 XX 10-DEC-2002; 2002US-0434778P.
 PR 10-MAR-2003; 2003US-0453135P.
 PR 30-APR-2003; 2003US-0466412P.
 PR 23-SEP-2003; 2003US-0504955P.
 XX
 XX (APPL-) APPLERA CORP.
 XX
 XX Cargill M, Devlin JJ, Iakubova O;
 XX WPI; 2004-533949/51.
 XX
 XX Identifying an individual who has an altered risk for developing
 PT myocardial infarction by detecting a single nucleotide polymorphism in
 PT the individual's nucleic acids.
 XX
 XX Claim 7; SEQ ID NO 9836; 145pp; English.
 XX
 CC The invention relates to a novel method for identifying an individual who
 CC has an altered risk for developing myocardial infarction. The method
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
 CC the nucleotide sequences given in the specification in the individual's
 CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC polynucleotide sequence represents a transcript-based context sequence
 CC which flanks an SNP found in a human myocardial infarction-associated
 CC gene of the invention. Note: This sequence was not shown in the
 CC specification. The sequence has come from an electronic sequence listing
 CC downloaded from the WIPO website.
 XX
 SQ Sequence 201 BP; 49 A; 61 C; 57 G; 33 T; 0 U; 1 Other;
 Query Match 78.2%; Score 17.2; DB 13; Length 201;
 Best Local Similarity 86.4%; Pred. No. 5.1e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CACCTTCCTTCATCCTCAGAC 22
 Db 89 CACCTTCCTTCCTCCTCAGAC 68
 RESULT 30
 ADQ48254/c
 ID ADQ48254 standard; DNA; 201 BP.
 XX
 XX ADQ48254;
 XX
 XX 18-NOV-2004 (first entry)
 XX
 XX Myocardial infarction-associated SNP flanking transcript, SEQ ID 9917.
 DE
 XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
 KW cardiant; gene therapy; human; ds.
 KW
 XX

OS Homo sapiens.
 XX WO2004058052-A2.
 XX
 XX 15-JUL-2004.
 XX
 XX 22-DEC-2003; 2003WO-US040978.
 XX
 XX 20-DEC-2002; 2002US-0434778P.
 PR 10-MAR-2003; 2003US-0453135P.
 PR 30-APR-2003; 2003US-0466412P.
 PR 23-SEP-2003; 2003US-0504955P.
 XX
 XX (APPL-) APPLERA CORP.
 XX
 XX Cargill M, Devlin JJ, Iakubova O;
 XX WPI; 2004-533949/51.
 XX
 XX Identifying an individual who has an altered risk for developing
 PT myocardial infarction by detecting a single nucleotide polymorphism in
 PT the individual's nucleic acids.
 XX
 XX Claim 7; SEQ ID NO 9917; 145pp; English.
 XX
 CC The invention relates to a novel method for identifying an individual who
 CC has an altered risk for developing myocardial infarction. The method
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
 CC the nucleotide sequences given in the specification in the individual's
 CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC polynucleotide sequence represents a transcript-based context sequence
 CC which flanks an SNP found in a human myocardial infarction-associated
 CC gene of the invention. Note: This sequence was not shown in the
 CC specification. The sequence has come from an electronic sequence listing
 CC downloaded from the WIPO website.
 XX
 SQ Sequence 201 BP; 49 A; 61 C; 57 G; 33 T; 0 U; 1 Other;
 Query Match 78.2%; Score 17.2; DB 13; Length 201;
 Best Local Similarity 86.4%; Pred. No. 5.1e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CACCTTCCTTCATCCTCAGAC 22
 Db 89 CACCTTCCTTCCTCCTCAGAC 68
 RESULT 31
 ADQ48102/c
 ID ADQ48102 standard; DNA; 201 BP.
 XX
 XX ADQ48102;
 XX
 XX 18-NOV-2004 (first entry)
 XX

DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 9765.
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX Homo sapiens.
XX WO2004058052-A2.
XX 15-JUL-2004.
XX 22-DEC-2003; 2003WO-US040978.
XX 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX (APPL-) APPLERA CORP.
XX Cargill M, Devlin JJ, Iakubova O;
PI WPI; 2004-533949/51.
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX Claim 7; SEQ ID NO 9765; 145pp; English.
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX Sequence 201 BP; 55 A; 54 C; 60 G; 31 T; 0 U; 1 Other;
SQ
Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CACCTTCCTTCCTCCTCAGAC 22
|||||||
Db 139 CACCTTCCTTCCTCCTCAGAC 118
RESULT 32.
ADQ48136/C
ID ADQ48136 standard; DNA; 201 BP.

XX ADQ48136;
AC 18-NOV-2004 (first entry)
XX
DT
XX
XX Myocardial infarction-associated SNP flanking transcript, SEQ ID 9799.
DE
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
KW
XX
XX Homo sapiens.
XX WO2004058052-A2.
XX 15-JUL-2004.
XX 22-DEC-2003; 2003WO-US040978.
XX 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX (APPL-) APPLERA CORP.
XX Cargill M, Devlin JJ, Iakubova O;
PI WPI; 2004-533949/51.
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX Claim 7; SEQ ID NO 9799; 145pp; English.
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX Sequence 201 BP; 55 A; 54 C; 60 G; 31 T; 0 U; 1 Other;
SQ
Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CACCTTCCTTCCTCCTCAGAC 22
|||||||
Db 139 CACCTTCCTTCCTCCTCAGAC 118

XX SQ Sequence 201 BP; 49 A; 61 C; 57 G; 33 T; 0 U; 1 Other;
Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 CACCTTCCTTCCTCCTCAGAC 22
Db 89 CACCTTCCTTCCTCCTCAGAC 68
RESULT 35
ADQ48231/c
ID ADQ48231 standard; DNA; 201 BP.
XX AC ADQ48231;
XX DT 18-NOV-2004 (first entry)
XX DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 9894.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human; ds.
XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX PI WPI; 2004-533949/51.
XX DR Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.
XX PS Claim 7; SEQ ID NO 9894; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
XX CC has an altered risk for developing myocardial infarction. The method
XX CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
XX CC the nucleotide sequences given in the specification in the individual's
XX CC nucleic acids, where the presence of the SNP is correlated with an
XX CC altered risk for myocardial infarction in the individual. The invention
XX CC further comprises: an isolated nucleic acid molecule comprising at least
XX CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
XX CC the specification or its complement and encoding any one of the amino
XX CC acid sequences given in the specification; an isolated polypeptide
XX CC comprising an amino acid sequence given in the specification; an antibody
XX CC that specifically binds to the polypeptide or its antigen-binding
XX CC fragment; an amplified polynucleotide containing an SNP given in the
XX CC specification and which is between about 16 and 1000 nucleotides in
XX CC length; a kit for detecting an SNP in a nucleic acid, comprising the
XX CC polynucleotide, a buffer and an enzyme; a method of detecting a SNP in a
XX CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
XX CC method for identifying an agent useful in treating or preventing
XX CC myocardial infarction. The novel detection method has cardiant activity.
XX CC The nucleic acids of the invention may be used in gene therapy. The
XX CC method is useful in identifying an individual who has an increased or
XX CC decreased risk for developing myocardial infarction and for preparing a
XX CC composition for treating or preventing myocardial infarction. This

CC Polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX SQ Sequence 201 BP; 55 A; 54 C; 60 G; 31 T; 0 U; 1 Other;
Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 CACCTTCCTTCCTCCTCAGAC 22
Db 139 CACCTTCCTTCCTCCTCAGAC 118
RESULT 36
ADQ48274/c
ID ADQ48274 standard; DNA; 201 BP.
XX AC ADQ48274;
XX DT 18-NOV-2004 (first entry)
XX DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 9937.
XX KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;
XX KW cardiant; gene therapy; human; ds.
XX OS Homo sapiens.
XX PN WO2004058052-A2.
XX PD 15-JUL-2004.
XX PF 22-DEC-2003; 2003WO-US040978.
XX PR 20-DEC-2002; 2002US-0434778P.
XX PR 10-MAR-2003; 2003US-0453135P.
XX PR 30-APR-2003; 2003US-0466412P.
XX PR 23-SEP-2003; 2003US-0504955P.
XX PA (APPL-) APPLERA CORP.
XX PI Cargill M, Devlin JJ, Iakoubova O;
XX PI WPI; 2004-533949/51.
XX DR Identifying an individual who has an altered risk for developing
XX PT myocardial infarction by detecting a single nucleotide polymorphism in
XX PT the individual's nucleic acids.
XX PS Claim 7; SEQ ID NO 9937; 145pp; English.
XX CC The invention relates to a novel method for identifying an individual who
XX CC has an altered risk for developing myocardial infarction. The method
XX CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
XX CC the nucleotide sequences given in the specification in the individual's
XX CC nucleic acids, where the presence of the SNP is correlated with an
XX CC altered risk for myocardial infarction in the individual. The invention
XX CC further comprises: an isolated nucleic acid molecule comprising at least
XX CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
XX CC the specification or its complement and encoding any one of the amino
XX CC acid sequences given in the specification; an isolated polypeptide
XX CC comprising an amino acid sequence given in the specification; an antibody
XX CC that specifically binds to the polypeptide or its antigen-binding
XX CC fragment; an amplified polynucleotide containing an SNP given in the
XX CC specification and which is between about 16 and 1000 nucleotides in
XX CC length; a kit for detecting an SNP in a nucleic acid, comprising the
XX CC polynucleotide, a buffer and an enzyme; a method of detecting a SNP in a
XX CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
XX CC method for identifying an agent useful in treating or preventing

CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX
SQ Sequence 201 BP; 48 A; 58 C; 60 G; 34 T; 0 U; 1 Other;

Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCTCAGAC 22
Db 71 CACCTTCCTTCCTCTCTCAGAC 50

RESULT 37
ADQ48129/c
ID ADQ48129 standard; DNA; 201 BP.
XX
AC ADQ48129;
XX
XX 18-NOV-2004 (first entry)
XX
XX Myocardial infarction-associated SNP flanking transcript, SEQ ID 9792.
DE
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
XX Homo sapiens.
OS
XX WO2004058052-A2.
PN
XX 15-JUL-2004.
PD
XX 22-DEC-2003; 2003WO-US040978.
PF
XX 20-DEC-2002; 2002US-0434778P.
PR
XX 10-MAR-2003; 2003US-0453135P.
PR
XX 30-APR-2003; 2003US-0466412P.
PR
XX 23-SEP-2003; 2003US-0504955P.
PR
XX (APPL-) APPLERA CORP.
PA
XX Cargill M, Devlin JJ, Iakoubova O;
PI
XX WPI; 2004-533949/51.
DR
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
XX Claim 7; SEQ ID NO 9792; 145pp; English.
PS
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the

CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.
XX
SQ Sequence 201 BP; 48 A; 58 C; 60 G; 34 T; 0 U; 1 Other;

Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCTCAGAC 22
Db 71 CACCTTCCTTCCTCTCTCAGAC 50

RESULT 38
ADQ48250/c
ID ADQ48250 standard; DNA; 201 BP.
XX
AC ADQ48250;
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Myocardial infarction-associated SNP flanking transcript, SEQ ID 9913.
DE
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
XX Homo sapiens.
OS
XX WO2004058052-A2.
PN
XX 15-JUL-2004.
PD
XX 22-DEC-2003; 2003WO-US040978.
PF
XX 20-DEC-2002; 2002US-0434778P.
PR
XX 10-MAR-2003; 2003US-0453135P.
PR
XX 30-APR-2003; 2003US-0466412P.
PR
XX 23-SEP-2003; 2003US-0504955P.
PR
XX (APPL-) APPLERA CORP.
PA
XX Cargill M, Devlin JJ, Iakoubova O;
PI
XX WPI; 2004-533949/51.
DR
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX
XX Claim 7; SEQ ID NO 9913; 145pp; English.
PS
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the

CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.

XX Sequence 201 BP; 50 A; 58 C; 57 G; 35 T; 0 U; 1 Other;

Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
|||||||
Db 65 CACCTTCCTTCTCTCCTCAGAC 44

RESULT 39
ADQ48333/c
ID ADQ48333 standard; DNA; 201 BP.
XX
AC ADQ48333;
XX
DT 18-NOV-2004 (first entry)
XX
DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 9996.
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
XX 15-JUL-2004.
PD
PF 22-DEC-2003; 2003WO-US040978.
XX
XX 20-DEC-2002; 2002US-0434778P.
PR
PR 10-MAR-2003; 2003US-0453135P.
PR
PR 30-APR-2003; 2003US-0466412P.
PR
PR 23-SEP-2003; 2003US-0504955P.
XX
XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakubova O;
XX
XX WPI; 2004-533949/51.
XX
XX Identifying an individual who has an altered risk for developing
XX myocardial infarction by detecting a single nucleotide polymorphism in
XX the individual's nucleic acids.

XX Claim 7; SEQ ID NO 9996; 145pp; English.
XX
XX The invention relates to a novel method for identifying an individual who
XX has an altered risk for developing myocardial infarction. The method
XX comprises detecting a single nucleotide polymorphism (SNP) in any one of

CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding
CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting a variant polypeptide; and a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC polynucleotide sequence represents a transcript-based context sequence
CC which flanks an SNP found in a human myocardial infarction-associated
CC gene of the invention. Note: This sequence was not shown in the
CC specification. The sequence has come from an electronic sequence listing
CC downloaded from the WIPO website.

XX Sequence 201 BP; 55 A; 54 C; 60 G; 31 T; 0 U; 1 Other;

Query Match 78.2%; Score 17.2; DB 13; Length 201;
Best Local Similarity 86.4%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
|||||||
Db 139 CACCTTCCTTCTCTCCTCAGAC 118

RESULT 40
ADQ48193/c
ID ADQ48193 standard; DNA; 201 BP.
XX
AC ADQ48193;
XX
DT 18-NOV-2004 (first entry)
XX
DE Myocardial infarction-associated SNP flanking transcript, SEQ ID 9856.
XX
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human; ds.
XX
OS Homo sapiens.
XX
PN WO2004058052-A2.
XX
XX 15-JUL-2004.
PD
PF 22-DEC-2003; 2003WO-US040978.
XX
XX 20-DEC-2002; 2002US-0434778P.
PR
PR 10-MAR-2003; 2003US-0453135P.
PR
PR 30-APR-2003; 2003US-0466412P.
PR
PR 23-SEP-2003; 2003US-0504955P.
XX
XX (APPL-) APPLERA CORP.

XX Cargill M, Devlin JJ, Iakubova O;
XX
XX WPI; 2004-533949/51.
XX
XX Identifying an individual who has an altered risk for developing
XX myocardial infarction by detecting a single nucleotide polymorphism in
XX the individual's nucleic acids.

PS Claim 7; SEQ ID NO 9856; 145pp; English.
 XX
 CC The invention relates to a novel method for identifying an individual who
 CC has an altered risk for developing myocardial infarction. The method
 CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
 CC the nucleotide sequences given in the specification in the individual's
 CC nucleic acids, where the presence of the SNP is correlated with an
 CC altered risk for myocardial infarction in the individual. The invention
 CC further comprises: an isolated nucleic acid molecule comprising at least
 CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC polynucleotide sequence represents a transcript-based context sequence
 CC which flanks an SNP found in a human myocardial infarction-associated
 CC gene of the invention. Note: This sequence was not shown in the
 CC specification. The sequence has come from an electronic sequence listing
 CC downloaded from the WIPO website.

XX
 SQ Sequence 201 BP; 48 A; 58 C; 60 G; 34 T; 0 U; 1 Other;
 Query Match 78.2%; Score 17.2; DB 13; Length 201;
 Best Local Similarity 86.4%; Pred. No. 5.1e+02;
 Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
 |||||
 Db 71 CACCTTCCTTCCTCCTCAGAC 50

Search completed: June 4, 2005, 07:28:53
 Job time : 168.757 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:48:59 ; Search time 48.9467 Seconds
(without alignments)
735.454 Million cell updates/sec

Title: US-09-674-277-22

Perfect score: 22

Sequence: 1 caccctccttcacacagac 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A COMB.seq:*

2: /cgn2_6/prodata/1/ina/5B COMB.seq:*

3: /cgn2_6/prodata/1/ina/6A COMB.seq:*

4: /cgn2_6/prodata/1/ina/6B COMB.seq:*

5: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*

6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19.4	88.2	601	4	US-09-949-016-148399
2	19.4	88.2	108169	4	US-09-949-016-12898
3	19.4	88.2	108169	4	US-09-949-016-15907
4	17.8	80.9	1001	3	US-09-841-638-452
5	17.8	80.9	1001	4	US-10-170-097-452
6	17.8	80.9	106418	4	US-09-949-016-13974
7	17.8	80.9	228851	4	US-09-949-016-13781
8	17.2	78.2	251	4	US-09-513-9990-2562
9	17.2	78.2	447	3	US-09-222-575-61
10	17.2	78.2	447	4	US-09-389-681-61
11	17.2	78.2	447	4	US-09-620-4059-61
12	17.2	78.2	447	4	US-09-339-338-61
13	17.2	78.2	447	4	US-09-433-826B-61
14	17.2	78.2	447	4	US-09-604-287A-61
15	17.2	78.2	447	4	US-09-285-480-61
16	17.2	78.2	447	4	US-09-834-759-61
17	17.2	78.2	447	4	US-09-590-751A-61
18	17.2	78.2	447	4	US-09-551-621-61
19	17.2	78.2	601	4	US-09-949-016-60446
20	17.2	78.2	601	4	US-09-949-016-76159
21	17.2	78.2	4078	4	US-09-016-434-1120
22	17.2	78.2	5926	4	US-09-917-254-41
23	17.2	78.2	10217	4	US-09-949-016-15556
24	17.2	78.2	11887	3	US-08-961-527-146
25	17.2	78.2	36227	4	US-09-949-016-13951
26	17.2	78.2	56694	4	US-09-949-016-12568
27	17.2	78.2	56702	4	US-09-949-016-15423

C 28	17.2	78.2	62804	3	US-09-800-960-3	Sequence 3, Appli
C 29	17.2	78.2	62804	4	US-10-096-960-3	Sequence 3, Appli
C 30	17.2	78.2	88758	4	US-09-949-016-13502	Sequence 15302, A
C 31	17.2	78.2	134987	4	US-09-949-016-15348	Sequence 15348, A
C 32	17.2	78.2	134987	4	US-09-949-016-15349	Sequence 15349, A
C 33	17.2	78.2	134987	4	US-09-949-016-15350	Sequence 15350, A
C 34	17.2	78.2	134987	4	US-09-949-016-15507	Sequence 15507, A
C 35	17.2	78.2	134987	4	US-09-949-016-15508	Sequence 15508, A
C 36	17.2	78.2	134987	4	US-09-949-016-15509	Sequence 15509, A
C 37	17.2	78.2	70563	4	US-09-949-016-16743	Sequence 16743, A
C 38	16.8	76.4	601	4	US-09-949-016-121965	Sequence 121965,
C 39	16.8	76.4	601	4	US-09-949-016-121966	Sequence 121966,
C 40	16.8	76.4	1389	2	US-08-023-980B-3	Sequence 3, Appli
C 41	16.8	76.4	1389	2	US-08-486-953A-3	Sequence 3, Appli
C 42	16.8	76.4	1389	4	US-08-204-052-3	Sequence 3, Appli
C 43	16.8	76.4	1396	6	5472691-1	Patent No. 5472691
C 44	16.8	76.4	1396	6	5472691-1	Patent No. 5472691
C 45	16.8	76.4	10079	2	US-08-476-866-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-09-949-016-148399
; Sequence 148399, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148399
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-148399

Query Match 88.2%; Score 19.4; DB 4; Length 601;
Best Local Similarity 95.2%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCTCAGA 21
| | | | | | | | | | | | | | | | | | | | | |
Db 12 CACCTTCCTTCCTCCTCAGA 32

RESULT 2

US-09-949-016-12898
; Sequence 12898, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12898
; LENGTH: 108169
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12898

Query Match      88.2%; Score 19.4; DB 4; Length 108169;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCATCCTCAGA 21
Db 38331 CACCTTCCTTCCTCCACCTCAGA 38351

RESULT 3
US-09-949-016-15907
; Sequence 15907, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15907
; LENGTH: 108169
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15907

Query Match      88.2%; Score 19.4; DB 4; Length 108169;
Best Local Similarity 95.2%; Pred. No. 28;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCATCCTCAGA 21
Db 38331 CACCTTCCTTCCTCCACCTCAGA 38351

RESULT 4
US-09-641-638-452
; Sequence 452, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
```

```
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 452
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-298-105 : polymorphic base G or A
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-298-105.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-298-105.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 586..605
; OTHER INFORMATION: upstream amplification primer, complement
; NAME/KEY: primer_bind
; LOCATION: 33..53
; OTHER INFORMATION: downstream amplification primer
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-298-105 potential probe
US-09-641-638-452

Query Match      80.9%; Score 17.8; DB 3; Length 1001;
Best Local Similarity 90.5%; Pred. No. 72;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCTCCATCCTCAGA 21
Db 759 CATCTTCCTTCCTTCCTCAGA 779

RESULT 5
US-10-170-097-452
; Sequence 452, Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 452
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-298-105 : polymorphic base G or A
; NAME/KEY: misc_binding
```

```
; LOCATION: 481..500
; OTHER INFORMATION: 12-298-105.mis1, potential
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-298-105.mis2, potential complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 586..605
; OTHER INFORMATION: upstream amplification primer, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 33..53
; OTHER INFORMATION: downstream amplification primer
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-298-105 potential probe
US-10-170-097-452

Query Match      80.9%; Score 17.8; DB 4; Length 1001;
Best Local Similarity 90.5%; Pred. No. 72;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGA 21
Db 759 CATCTTCCTTCCCTCCTCAGA 779

RESULT 6
US-09-949-016-13974/c
; Sequence 13974, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13974
; LENGTH: 106418
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(106418)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13974

Query Match      80.9%; Score 17.8; DB 4; Length 106418;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGA 21
Db 100670 CAGCTTCCTTCCAGCCTCAGA 100650

RESULT 7
US-09-949-016-13781/c
; Sequence 13781, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13974
; LENGTH: 106418
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(106418)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13974

Query Match      80.9%; Score 17.8; DB 4; Length 106418;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGA 21
Db 100670 CAGCTTCCTTCCAGCCTCAGA 100650

RESULT 8
US-09-513-999C-2562/c
; Sequence 2562, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2 REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 2562
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 67..249
US-09-513-999C-2562

Query Match      78.2%; Score 17.2; DB 4; Length 251;
Best Local Similarity 86.4%; Pred. No. 1.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
Db 132 CACCTTCCTTCCATGCTCCGAC 111

RESULT 9
US-09-222-575-61
; Sequence 61, Application US/09222575
; Patent No. 6387697
; GENERAL INFORMATION:
; APPLICANT: Yugu, Jiang
```

```
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: Compositions for the Treatment and Diagnosis of Breast Cancer
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 210121.470
; CURRENT APPLICATION NUMBER: US/09/222,575
; CURRENT FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-222-575-61

Query Match      78.2%; Score 17.2; DB 3; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  CACCTTCCTTCCATCCTCAGAC 22
      ||||| ||||| ||||| |||||
Db      345  CACCCTCCTTCCATGCTCCGAC 366

RESULT 10
US-09-389-681-61
; Sequence 61, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389-681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-389-681-61

Query Match      78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  CACCTTCCTTCCATCCTCAGAC 22
      ||||| ||||| ||||| |||||
Db      345  CACCCTCCTTCCATGCTCCGAC 366

RESULT 11
US-09-620-405B-61
; Sequence 61, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-620-405B-61

Query Match      78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  CACCTTCCTTCCATCCTCAGAC 22
      ||||| ||||| ||||| |||||
Db      345  CACCCTCCTTCCATGCTCCGAC 366

RESULT 12
US-09-339-338-61
; Sequence 61, Application US/09339338A
; Patent No. 6573368
; GENERAL INFORMATION:
; APPLICANT: Yuqui, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C2
; CURRENT APPLICATION NUMBER: US/09/339,338A
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 315
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-339-338-61

Query Match      78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  CACCTTCCTTCCATCCTCAGAC 22
      ||||| ||||| ||||| |||||
Db      345  CACCCTCCTTCCATGCTCCGAC 366

RESULT 13
US-09-433-826B-61
; Sequence 61, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-433-826B-61

Query Match      78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```



```
QY 1 CACCTTCCTTCCATCCTCAGAC 22
    ||| ||| ||| ||| ||| ||| ||| |||
Db 345 CACCTTCCTTCCATGCTCCGAC 366

RESULT 14
US-09-604-287A-61
; Sequence 61, Application US/09604287A
; Patent No. 6586572
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C7
; CURRENT APPLICATION NUMBER: US/09/604,287A
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 489
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-604-287A-61

Query Match 78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
    ||| ||| ||| ||| ||| ||| ||| |||
Db 345 CACCTTCCTTCCATGCTCCGAC 366

RESULT 15
US-09-285-480-61
; Sequence 61, Application US/09285480
; Patent No. 6590076
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C1
; CURRENT APPLICATION NUMBER: US/09/285,480
; CURRENT FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-285-480-61

Query Match 78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
    ||| ||| ||| ||| ||| ||| ||| |||
Db 345 CACCTTCCTTCCATGCTCCGAC 366

RESULT 16
US-09-834-759-61
; Sequence 61, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-834-759-61

Query Match 78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
    ||| ||| ||| ||| ||| ||| ||| |||
Db 345 CACCTTCCTTCCATGCTCCGAC 366

RESULT 17
US-09-590-751A-61
; Sequence 61, Application US/09590751A
; Patent No. 6756477
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C6
; CURRENT APPLICATION NUMBER: US/09/590,751A
; CURRENT FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-590-751A-61

Query Match 78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
    ||| ||| ||| ||| ||| ||| ||| |||
Db 345 CACCTTCCTTCCATGCTCCGAC 366

RESULT 18
US-09-551-621-61
; Sequence 61, Application US/09551621
; Patent No. 6825175
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
```

```
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; FILE REFERENCE: 210121.470C5
; CURRENT APPLICATION NUMBER: US/09/551,621
; CURRENT FILING DATE: 2000-04-17
; NUMBER OF SEQ ID NOS: 479
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 61
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-551-621-61

Query Match          78.2%; Score 17.2; DB 4; Length 447;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
Db 345 CACCTTCCTTCATGCTCCGAC 366

RESULT 19
US-09-949-016-60446/c
; Sequence 60446, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 60446
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-60446

Query Match          78.2%; Score 17.2; DB 4; Length 601;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
Db 405 CTCTGCTTCAGGCTCAGAC 384

RESULT 20
US-09-949-016-76159/c
; Sequence 76159, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76159
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-76159

Query Match          78.2%; Score 17.2; DB 4; Length 601;
Best Local Similarity 86.4%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
Db 443 CACCTTCCTTTAATCCTCAGAC 422

RESULT 21
US-09-016-434-1120/c
; Sequence 1120, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSER: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1120:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4078 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1377819
US-09-016-434-1120

Query Match          78.2%; Score 17.2; DB 4; Length 4078;
Best Local Similarity 86.4%; Pred. No. 1.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
Db 2954 CACCTTCCTTCCTCCTCAGAC 2933
```

```
RESULT 22
US-09-917-254-41/c
; Sequence 41, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
; LENGTH: 5926
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-09-917-254-41
Query Match 78.2%; Score 17.2; DB 4; Length 5926;
Best Local Similarity 86.4%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
||||| ||||| ||||| |||||
Db 2954 CACCTTCCTTCCTCCTCAGAC 2933

RESULT 23
US-09-949-016-15556
; Sequence 15556, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15556
; LENGTH: 10217
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15556
Query Match 78.2%; Score 17.2; DB 4; Length 10217;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
||||| ||||| ||||| |||||
Db 163 CACCTTCCTTCCTCCTCAGAC 184

RESULT 24
US-08-961-527-146/c
; Sequence 146, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 146:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11887 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-146
Query Match 78.2%; Score 17.2; DB 3; Length 11887;
Best Local Similarity 86.4%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
||||| ||||| ||||| |||||
Db 3655 CGCCTTTTTCATCCTCAGAC 3634

RESULT 25
US-09-949-016-13951/c
; Sequence 13951, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13951
; LENGTH: 36227
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13951
Query Match 78.2%; Score 17.2; DB 4; Length 36227;
Best Local Similarity 86.4%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
```

Db	9559	CACCTTTCTTAACTCCTCAGAC	9538		
RESULT 26					
US-09-949-016-12568/c					
; Sequence 12568, Application US/09949016					
; Patent No. 6812339					
; GENERAL INFORMATION:					
; APPLICANT: VENTER, J. Craig et al.					
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF					
; FILE REFERENCE: CL001307					
; CURRENT APPLICATION NUMBER: US/09/949,016					
; CURRENT FILING DATE: 2000-04-14					
; PRIOR APPLICATION NUMBER: 60/241,755					
; PRIOR FILING DATE: 2000-10-20					
; PRIOR APPLICATION NUMBER: 60/237,768					
; PRIOR FILING DATE: 2000-10-03					
; PRIOR APPLICATION NUMBER: 60/231,498					
; PRIOR FILING DATE: 2000-09-08					
; NUMBER OF SEQ ID NOS: 207012					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 12568					
; LENGTH: 56694					
; TYPE: DNA					
; ORGANISM: Human					
US-09-949-016-12568					
Query Match 78.2%; Score 17.2; DB 4; Length 56694;					
Best Local Similarity 86.4%; Pred. No. 2.5e+02;					
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
QY	1	CACCTTCCTTCCATCCTCAGAC	22		
Db	13841	CATCTTCTTCCATCCTCAGAC	13820	CATCTTCTTCCATCCTCAGAC	
RESULT 27					
US-09-949-016-15423/c					
; Sequence 15423, Application US/09949016					
; Patent No. 6812339					
; GENERAL INFORMATION:					
; APPLICANT: VENTER, J. Craig et al.					
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF					
; FILE REFERENCE: CL001307					
; CURRENT APPLICATION NUMBER: US/09/949,016					
; CURRENT FILING DATE: 2000-04-14					
; PRIOR APPLICATION NUMBER: 60/241,755					
; PRIOR FILING DATE: 2000-10-20					
; PRIOR APPLICATION NUMBER: 60/237,768					
; PRIOR FILING DATE: 2000-10-03					
; PRIOR APPLICATION NUMBER: 60/231,498					
; PRIOR FILING DATE: 2000-09-08					
; NUMBER OF SEQ ID NOS: 207012					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 15423					
; LENGTH: 56702					
; TYPE: DNA					
; ORGANISM: Human					
US-09-949-016-15423					
Query Match 78.2%; Score 17.2; DB 4; Length 56702;					
Best Local Similarity 86.4%; Pred. No. 2.5e+02;					
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
QY	1	CACCTTCCTTCCATCCTCAGAC	22		
Db	13841	CATCTTCTTCCATCCTCAGAC	13820	CATCTTCTTCCATCCTCAGAC	
RESULT 28					
US-09-949-016-13502					
; Sequence 13502, Application US/09949016					
; Patent No. 6812339					
; GENERAL INFORMATION:					
; APPLICANT: VENTER, J. Craig et al.					

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13502
; LENGTH: 88758
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13502

Query Match          78.2%; Score 17.2; DB 4; Length 88758;
Best Local Similarity 86.4%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
Db 57904 CTCCTGCTTCCAGCTCAGAC 57925

RESULT 31
US-09-949-016-15348
; Sequence 15348, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15348
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15348

Query Match          78.2%; Score 17.2; DB 4; Length 134987;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
Db 28631 CTCCTTCCTTCCCTCCCGAC 28652

RESULT 32
US-09-949-016-15349
; Sequence 15349, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15349
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15349

Query Match          78.2%; Score 17.2; DB 4; Length 134987;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
Db 28631 CTCCTTCCTTCCCTCCCGAC 28652

RESULT 33
US-09-949-016-15350
; Sequence 15350, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15350
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15350

Query Match          78.2%; Score 17.2; DB 4; Length 134987;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
Db 28631 CTCCTTCCTTCCCTCCCGAC 28652

RESULT 34
US-09-949-016-15507
; Sequence 15507, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15507
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15507

Query Match          78.2%; Score 17.2; DB 4; Length 134987;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   |||||
Db 28631 CTCCTTCCTTCCCTCCCGAC 28652

RESULT 35
US-09-949-016-15508
; Sequence 15508, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15508
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15508

Query Match          78.2%; Score 17.2; DB 4; Length 134987;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   |||||
Db 28631 CTCCTTCCTTCCCTCCCGAC 28652

RESULT 36
US-09-949-016-15509
; Sequence 15509, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15509
; LENGTH: 134987

Query Match          78.2%; Score 17.2; DB 4; Length 134987;
Best Local Similarity 86.4%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   |||||
Db 28631 CTCCTTCCTTCCCTCCCGAC 28652

RESULT 37
US-09-949-016-16743
; Sequence 16743, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16743
; LENGTH: 70563
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16743

Query Match          77.3%; Score 17; DB 4; Length 70563;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 TTCCTTCCTTCCATCCTCAGA 21
   |||||
Db 20007 TTCCTTCCTTCCATCCTCAGA 20023

RESULT 38
US-09-949-016-121965
; Sequence 121965, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121965
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-121965

Query Match          76.4%; Score 16.8; DB 4; Length 601;
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 07:08:50 ; Search time 209.716 Seconds
(without alignments)
644.888 Million cell updates/sec

Title: US-09-674-277-22

Perfect score: 22

Sequence: 1 caccctccctccatccacag 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	90.9	341	9	US-09-983-965-1524
C 2	18.8	85.5	186391	13	US-10-087-192-136
C 3	17.8	80.9	333	18	US-10-437-963-95765
C 4	17.8	80.9	592	13	US-10-040-739-916
C 5	17.8	80.9	1001	17	US-10-170-097-452
C 6	17.8	80.9	1001	19	US-10-926-668-452
C 7	17.8	80.9	1251	18	US-10-437-963-40466
C 8	17.8	80.9	4997	17	US-10-260-238-1167
C 9	17.8	80.9	238910	18	US-10-719-993-6801
C 10	17.4	79.1	650	13	US-10-027-632-204217
C 11	17.4	79.1	650	17	US-10-027-632-204217

C 12	17.4	79.1	1530	9	US-09-938-842A-1553	Sequence 1553, Ap
C 13	17.4	79.1	1530	11	US-09-938-842A-1553	Sequence 1553, Ap
C 14	17.4	79.1	32572	18	US-10-741-601-5701	Sequence 5701, Ap
C 15	17.4	79.1	32572	19	US-10-741-600-17802	Sequence 17802, A
C 16	17.4	79.1	80399	18	US-10-719-993-6831	Sequence 6831, Ap
C 17	17.4	79.1	165221	13	US-10-087-192-1015	Sequence 1015, Ap
C 18	17.4	79.1	78.2	201	US-10-741-600-9763	Sequence 9763, Ap
C 19	17.2	78.2	78.2	201	US-10-741-600-9765	Sequence 9765, Ap
C 20	17.2	78.2	78.2	201	US-10-741-600-9766	Sequence 9766, Ap
C 21	17.2	78.2	78.2	201	US-10-741-600-9771	Sequence 9771, Ap
C 22	17.2	78.2	78.2	201	US-10-741-600-9789	Sequence 9789, Ap
C 23	17.2	78.2	78.2	201	US-10-741-600-9792	Sequence 9792, Ap
C 24	17.2	78.2	78.2	201	US-10-741-600-9794	Sequence 9794, Ap
C 25	17.2	78.2	78.2	201	US-10-741-600-9797	Sequence 9797, Ap
C 26	17.2	78.2	78.2	201	US-10-741-600-9799	Sequence 9799, Ap
C 27	17.2	78.2	78.2	201	US-10-741-600-9800	Sequence 9800, Ap
C 28	17.2	78.2	78.2	201	US-10-741-600-9804	Sequence 9804, Ap
C 29	17.2	78.2	78.2	201	US-10-741-600-9815	Sequence 9815, Ap
C 30	17.2	78.2	78.2	201	US-10-741-600-9817	Sequence 9817, Ap
C 31	17.2	78.2	78.2	201	US-10-741-600-9818	Sequence 9818, Ap
C 32	17.2	78.2	78.2	201	US-10-741-600-9836	Sequence 9836, Ap
C 33	17.2	78.2	78.2	201	US-10-741-600-9838	Sequence 9838, Ap
C 34	17.2	78.2	78.2	201	US-10-741-600-9839	Sequence 9839, Ap
C 35	17.2	78.2	78.2	201	US-10-741-600-9843	Sequence 9843, Ap
C 36	17.2	78.2	78.2	201	US-10-741-600-9854	Sequence 9854, Ap
C 37	17.2	78.2	78.2	201	US-10-741-600-9856	Sequence 9856, Ap
C 38	17.2	78.2	78.2	201	US-10-741-600-9857	Sequence 9857, Ap
C 39	17.2	78.2	78.2	201	US-10-741-600-9891	Sequence 9891, Ap
C 40	17.2	78.2	78.2	201	US-10-741-600-9894	Sequence 9894, Ap
C 41	17.2	78.2	78.2	201	US-10-741-600-9895	Sequence 9895, Ap
C 42	17.2	78.2	78.2	201	US-10-741-600-9897	Sequence 9897, Ap
C 43	17.2	78.2	78.2	201	US-10-741-600-9910	Sequence 9910, Ap
C 44	17.2	78.2	78.2	201	US-10-741-600-9912	Sequence 9912, Ap
C 45	17.2	78.2	78.2	201	US-10-741-600-9913	Sequence 9913, Ap

ALIGNMENTS

RESULT 1

US-09-983-965-1524/c
; Sequence 1524, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 1524
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (297),(300)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 37-LIB2809-032-Q1-E1-B2
US-09-983-965-1524

Query Match 90.9%; Score 20; DB 9; Length 341;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTTTCCTTCATCCTCAGAC 22
| | | | | | | | | | | | | | |
Db 265 CTTTCCTTCATCCTCAGAC 246

RESULT 2
US-10-087-192-136
; Sequence 136, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 136
; LENGTH: 186391
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(186391)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-136

Query Match 85.5%; Score 18.8; DB 13; Length 186391;
Best Local Similarity 90.9%; Pred. No. 66;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
| | | | | | | | | | | | | | |
Db 21054 CACCTTCCTCCCATCCTCAGAC 21075

RESULT 3
US-10-437-963-95765
; Sequence 95765, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95765
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_93927C.1
US-10-437-963-95765

Query Match 80.9%; Score 17.8; DB 18; Length 393;
Best Local Similarity 90.5%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGA 21
| | | | | | | | | | | | | | |
Db 241 CACCTTCCTTCCTCCTCAGA 261

RESULT 4
US-10-040-739-916
; Sequence 916, Application US/10040739
; Publication No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaValle, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 916:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 592 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 916:
US-10-040-739-916

Query Match 80.9%; Score 17.8; DB 13; Length 592;
Best Local Similarity 90.5%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGA 21
| | | | | | | | | | | | | | |
Db 108 CAGCTTCCTTCATCCTGAGA 128

RESULT 5
US-10-170-097-452
; Sequence 452, Application US/10170097
; Publication No. US20030228582A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING

```
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GEN-T14XC2D1
; CURRENT APPLICATION NUMBER: US/10/170,097
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 452
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-298-105 : polymorphic base G or A
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-298-105.mis1, potential
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 502..521
; OTHER INFORMATION: 12-298-105.mis2, potential complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 586..605
; OTHER INFORMATION: upstream amplification primer, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 33..53
; OTHER INFORMATION: downstream amplification primer
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 499..513
; OTHER INFORMATION: 12-298-105 potential probe
; US-10-170-097-452

Query Match      80.9%; Score 17.8; DB 17; Length 1001;
Best Local Similarity 90.5%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CACCTTCCTTCCTCCTCCTCAGA 21
Db      759 CATCTTCCTTCCTTCCTCCTCAGA 779

RESULT 6
US-10-926-684-452
; Sequence 452, Application US/10926684
; Publication No. US20050014190A1
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marica
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS
; TITLE OF INVENTION: CARRYING
; FILE REFERENCE: GEN-T14XC2D1
; CURRENT APPLICATION NUMBER: US/10/926,684
; CURRENT FILING DATE: 2004-08-25
; PRIOR APPLICATION NUMBER: US/10/170,097
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
```

```
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 452
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-298-105 : polymorphic base G or A
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-298-105.mis1, potential
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-298-105.mis2, potential complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 586..605
; OTHER INFORMATION: upstream amplification primer, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 33..53
; OTHER INFORMATION: downstream amplification primer
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 499..513
; OTHER INFORMATION: 12-298-105 potential probe
; US-10-926-684-452

Query Match      80.9%; Score 17.8; DB 19; Length 1001;
Best Local Similarity 90.5%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 CACCTTCCTTCCTCCTCCTCAGA 21
Db      759 CATCTTCCTTCCTTCCTCCTCAGA 779

RESULT 7
US-10-437-963-40466/c
; Sequence 40466, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 40466
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Oryza sativa
```

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_43906C.1
US-10-437-963-40466

Query Match 80.9%; Score 17.8; DB 18; Length 1251;
Best Local Similarity 90.5%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 2;

QY 2 ACCTTCCTTCCATCCTCAGAC 22

Db 1142 ACCTTCCTTCCATCCTCGGAC 1122

RESULT 8

US-10-260-238-1167
Sequence 1167, Application US/10260238
Publication No. US20040016025A1
GENERAL INFORMATION:
APPLICANT: Budworth, Paul R.
APPLICANT: Moughamer, Todd G.
APPLICANT: Briggs, Steven P.
APPLICANT: Cooper, Bret
APPLICANT: Glazebrook, Jane
APPLICANT: Goff, Stephen A.
APPLICANT: Katagiri, FumiYaki
APPLICANT: Kreps, Joel
APPLICANT: Provart, Nicholas
APPLICANT: Ricke, Darrell
APPLICANT: Zhu, Tong
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
FILE REFERENCE: 60111-NP
CURRENT APPLICATION NUMBER: US/10/260,238
CURRENT FILING DATE: 2002-09-26
PRIOR APPLICATION NUMBER: US 60/325,448
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/325,277
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: US 60/370,620
PRIOR FILING DATE: 2002-04-04
NUMBER OF SEQ ID NOS: 6077
SEQ ID NO 1167
LENGTH: 4997
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: N region
LOCATION: (3713)..(3713)
OTHER INFORMATION: n = any nucleotide
FEATURE:
NAME/KEY: N region
LOCATION: (3765)..(3765)
OTHER INFORMATION: n = any nucleotide
FEATURE:
NAME/KEY: N region
LOCATION: (4190)..(4190)
OTHER INFORMATION: n = any nucleotide
FEATURE:
NAME/KEY: N region
LOCATION: (4195)..(4195)
OTHER INFORMATION: n = any nucleotide
FEATURE:
NAME/KEY: N region
LOCATION: (4220)..(4220)
OTHER INFORMATION: n = any nucleotide
US-10-260-238-1167

Query Match 80.9%; Score 17.8; DB 17; Length 4997;
Best Local Similarity 90.5%; Pred. No. 2e+02; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 2;

QY 1 CACCTTCCTTCCATCCTCAGA 21

Db 4666 CACCTTCCTTCCGTCATCAGA 4686

RESULT 9

US-10-719-993-6801/c
Sequence 6801, Application US/10719993
Publication No. US20040265849A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001496
CURRENT APPLICATION NUMBER: US/10/719,993
CURRENT FILING DATE: 2003-11-24
NUMBER OF SEQ ID NOS: 55342
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6801
LENGTH: 238910
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(238910)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-719-993-6801

Query Match 80.9%; Score 17.8; DB 18; Length 238910;
Best Local Similarity 90.5%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGA 21

Db 226924 CAGCTTCCTTCCAGCCTCAGA 226904

RESULT 10

US-10-027-632-204217/c
Sequence 204217, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
TITLE OF INVENTION: Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 204217
LENGTH: 650
TYPE: DNA
ORGANISM: Human
US-10-027-632-204217

Query Match 79.1%; Score 17.4; DB 13; Length 650;
Best Local Similarity 94.7%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCA 19

|||

Db 453 CAACCTTCCTCCATCCTCA 435

RESULT 11

US-10-027-632-204217/c

; Sequence 204217, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 204217

; LENGTH: 650

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-204217

Query Match 79.1%; Score 17.4; DB 17; Length 650;

Best Local Similarity 94.7%; Pred. No. 3.1e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAACCTTCCTCCATCCTCA 19

Db 453 CAACCTTCCTCCATCCTCA 435

RESULT 12

US-09-938-842A-1553/c

; Sequence 1553, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1553

; LENGTH: 1530

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-1553

Query Match 79.1%; Score 17.4; DB 9; Length 1530;

Best Local Similarity 94.7%; Pred. No. 3.1e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Best Local Similarity 94.7%; Pred. No. 3e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTTCCTTCCTCCATCCTCA 21

Db 711 CCTTCCTTCCTCCATCCTCA 693

RESULT 13

US-09-938-842A-1553/c

; Sequence 1553, Application US/09938842A

; Publication No. US20040009476A9

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 1553

; LENGTH: 1530

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-1553

Query Match 79.1%; Score 17.4; DB 11; Length 1530;

Best Local Similarity 94.7%; Pred. No. 3e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTTCCTTCCTCCATCCTCA 21

Db 711 CCTTCCTTCCTCCATCCTCA 693

RESULT 14

US-10-741-601-5701/c

; Sequence 5701, Application US/10741601

; Publication No. US20040166519A1

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; FILE REFERENCE: CL001500

; CURRENT APPLICATION NUMBER: US/10/741,601

; CURRENT FILING DATE: 2003-12-22

; NUMBER OF SEQ ID NOS: 26415

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5701

; LENGTH: 32572

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-741-601-5701

Query Match 79.1%; Score 17.4; DB 18; Length 32572;

Best Local Similarity 94.7%; Pred. No. 2.9e+02;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTTTCCTTCCTCCATCCTCAGAC 22

Db 28857 CTTTCCTTCCTCCATCCTCAGAC 28839

RESULT 15

```
US-10-741-600-17802/c
; Sequence 17802, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17802
; LENGTH: 32572
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-17802

Query Match      79.1%; Score 17.4; DB 19; Length 32572;
Best Local Similarity 94.7%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 CTTCTTCCATCCTCAGAC 22
      |||||
Db      28857 CTTCTTCCATCCTCAGAC 28839

RESULT 16
US-10-719-993-6831
; Sequence 6831, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6831
; LENGTH: 80399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-6831

Query Match      79.1%; Score 17.4; DB 18; Length 80399;
Best Local Similarity 94.7%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 CCTTCTTCCATCCTCAGA 21
      |||||
Db      25668 CCTTCTTCCATCCTCAGA 25686

RESULT 17
US-10-087-192-1015/c
; Sequence 1015, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1015
; LENGTH: 165221
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(165221)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1015

Query Match      79.1%; Score 17.4; DB 13; Length 165221;
Best Local Similarity 94.7%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 CTTCTTCCATCCTCAGAC 22
      |||||
Db      48021 CTTCTTCCATCCTCAGAC 48003

RESULT 18
US-10-741-600-9763/c
; Sequence 9763, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9763
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9763

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 CACCTTCTTCCATCCTCAGAC 22
      |||||
Db      89 CACCTTCTTCTCTCTCCTCAGAC 68

RESULT 19
US-10-741-600-9765/c
; Sequence 9765, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9765
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9765

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 CACCTTCTTCTCTCTCCTCAGAC 22
      |||||
Db      139 CACCTTCTTCTCTCTCCTCAGAC 118
```

```
RESULT 20
US-10-741-600-9766/c
; Sequence 9766, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9766
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9766

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
   ||||| ||||| ||||| |||||
DB 158 CACCTTCCTTCCTCCTCAGAC 137

RESULT 21
US-10-741-600-9771/c
; Sequence 9771, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9771
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9771

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
   ||||| ||||| ||||| |||||
DB 158 CACCTTCCTTCCTCCTCAGAC 137

RESULT 22
US-10-741-600-9789/c
; Sequence 9789, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9789
; LENGTH: 201
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9789

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
   ||||| ||||| ||||| |||||
DB 194 CACCTTCCTTCCTCCTCAGAC 173

RESULT 23
US-10-741-600-9792/c
; Sequence 9792, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9792
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9792

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
   ||||| ||||| ||||| |||||
DB 71 CACCTTCCTTCCTCCTCAGAC 50

RESULT 24
US-10-741-600-9794/c
; Sequence 9794, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9794
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9794

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
   ||||| ||||| ||||| |||||
DB 65 CACCTTCCTTCCTCCTCAGAC 44

RESULT 25
US-10-741-600-9797/c
; Sequence 9797, Application US/10741600
; Publication No. US20050026169A1
```

```
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9797
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9797

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   ||||| | | | | | | | |
Db 89 CACCTTCCTCTCTCTCAGAC 68

RESULT 26
US-10-741-600-9799/c
; Sequence 9799, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9799
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9799

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   ||||| | | | | | | | |
Db 139 CACCTTCCTCTCTCTCAGAC 118

RESULT 27
US-10-741-600-9800/c
; Sequence 9800, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9800
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9800

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
```

```
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   ||||| | | | | | | | |
Db 158 CACCTTCCTCTCTCTCAGAC 137

RESULT 28
US-10-741-600-9804/c
; Sequence 9804, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9804
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9804

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   ||||| | | | | | | | |
Db 179 CACCTTCCTCTCTCTCAGAC 158

RESULT 29
US-10-741-600-9815/c
; Sequence 9815, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9815
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9815

Query Match      78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
   ||||| | | | | | | | |
Db 194 CACCTTCCTCTCTCTCAGAC 173

RESULT 30
US-10-741-600-9817/c
; Sequence 9817, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
```


; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9817
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9817

Query Match 78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
|||||||
DB 71 CACCTTCCTTCTCTCTCAGAC 50

RESULT 31
US-10-741-600-9818/c
; Sequence 9818, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9818
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9818

Query Match 78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
|||||||
DB 65 CACCTTCCTTCTCTCTCAGAC 44

RESULT 32
US-10-741-600-9836/c
; Sequence 9836, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9836
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9836

Query Match 78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
|||||||
DB 89 CACCTTCCTTCTCTCTCAGAC 68

RESULT 33
US-10-741-600-9838/c
; Sequence 9838, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9838
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9838

Query Match 78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
|||||||
DB 139 CACCTTCCTTCTCTCTCAGAC 118

RESULT 34
US-10-741-600-9839/c
; Sequence 9839, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9839
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-9839

Query Match 78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCAGAC 22
|||||||
DB 158 CACCTTCCTTCTCTCTCAGAC 137

RESULT 35
US-10-741-600-9843/c
; Sequence 9843, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9843
; LENGTH: 201
; TYPE: DNA

APPLICANT	GENETIC POLYMORPHISMS ASSOCIATED WITH	MYOCARDIAL INFARCTION	METHODS OF DETECTION AND USES THEREOF
US-10-741-600-9843	US-10-741-600-9843	US-10-741-600-9843	US-10-741-600-9843
Query Match	78.2%; Score 17.2; DB 19; Length 201;	78.2%; Score 17.2; DB 19; Length 201;	78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity	86.4%; Pred. No. 3.8e+02;	86.4%; Pred. No. 3.8e+02;	86.4%; Pred. No. 3.8e+02;
Matches	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1 CACCTTCCTCCATCCTCAGAC 22	1 CACCTTCCTCCATCCTCAGAC 22	1 CACCTTCCTCCATCCTCAGAC 22
Db	179 CACCTTCCTCTCTCTCCTCAGAC 158	65 CACCTTCCTCTCTCTCCTCAGAC 44	89 CACCTTCCTCTCTCTCCTCAGAC 68
RESULT 36	US-10-741-600-9854/c	US-10-741-600-9857/c	US-10-741-600-9894/c
Sequence 9854, Application US/10741600	Sequence 9854, Application US/10741600	Sequence 9894, Application US/10741600	Sequence 9894, Application US/10741600
Publication No. US20050026169A1	Publication No. US20050026169A1	Publication No. US20050026169A1	Publication No. US20050026169A1
GENERAL INFORMATION:	GENERAL INFORMATION:	GENERAL INFORMATION:	GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.	APPLICANT: CARGILL, Michele et al.	APPLICANT: CARGILL, Michele et al.	APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH	TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH	TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH	TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001499	FILE REFERENCE: CL001499	FILE REFERENCE: CL001499	FILE REFERENCE: CL001499
CURRENT APPLICATION NUMBER: US/10/741,600	CURRENT APPLICATION NUMBER: US/10/741,600	CURRENT APPLICATION NUMBER: US/10/741,600	CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22	CURRENT FILING DATE: 2003-12-22	CURRENT FILING DATE: 2003-12-22	CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997	NUMBER OF SEQ ID NOS: 73997	NUMBER OF SEQ ID NOS: 73997	NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0	SOFTWARE: FastSeq for Windows Version 4.0	SOFTWARE: FastSeq for Windows Version 4.0	SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9854	SEQ ID NO 9854	SEQ ID NO 9891	SEQ ID NO 9894
LENGTH: 201	LENGTH: 201	LENGTH: 201	LENGTH: 201
TYPE: DNA	TYPE: DNA	TYPE: DNA	TYPE: DNA
ORGANISM: Homo sapiens	ORGANISM: Homo sapiens	ORGANISM: Homo sapiens	ORGANISM: Homo sapiens
US-10-741-600-9854	US-10-741-600-9854	US-10-741-600-9891	US-10-741-600-9894
Query Match	78.2%; Score 17.2; DB 19; Length 201;	78.2%; Score 17.2; DB 19; Length 201;	78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity	86.4%; Pred. No. 3.8e+02;	86.4%; Pred. No. 3.8e+02;	86.4%; Pred. No. 3.8e+02;
Matches	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1 CACCTTCCTCCATCCTCAGAC 22	1 CACCTTCCTCCATCCTCAGAC 22	1 CACCTTCCTCCATCCTCAGAC 22
Db	194 CACCTTCCTCTCTCTCCTCAGAC 173	65 CACCTTCCTCTCTCTCCTCAGAC 44	89 CACCTTCCTCTCTCTCCTCAGAC 68
RESULT 37	US-10-741-600-9856/c	US-10-741-600-9891/c	US-10-741-600-9894/c
Sequence 9856, Application US/10741600	Sequence 9856, Application US/10741600	Sequence 9891, Application US/10741600	Sequence 9894, Application US/10741600
Publication No. US20050026169A1	Publication No. US20050026169A1	Publication No. US20050026169A1	Publication No. US20050026169A1
GENERAL INFORMATION:	GENERAL INFORMATION:	GENERAL INFORMATION:	GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.	APPLICANT: CARGILL, Michele et al.	APPLICANT: CARGILL, Michele et al.	APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH	TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH	TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH	TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCE: CL001499	FILE REFERENCE: CL001499	FILE REFERENCE: CL001499	FILE REFERENCE: CL001499
CURRENT APPLICATION NUMBER: US/10/741,600	CURRENT APPLICATION NUMBER: US/10/741,600	CURRENT APPLICATION NUMBER: US/10/741,600	CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22	CURRENT FILING DATE: 2003-12-22	CURRENT FILING DATE: 2003-12-22	CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997	NUMBER OF SEQ ID NOS: 73997	NUMBER OF SEQ ID NOS: 73997	NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0	SOFTWARE: FastSeq for Windows Version 4.0	SOFTWARE: FastSeq for Windows Version 4.0	SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9856	SEQ ID NO 9856	SEQ ID NO 9891	SEQ ID NO 9894
LENGTH: 201	LENGTH: 201	LENGTH: 201	LENGTH: 201
TYPE: DNA	TYPE: DNA	TYPE: DNA	TYPE: DNA
ORGANISM: Homo sapiens	ORGANISM: Homo sapiens	ORGANISM: Homo sapiens	ORGANISM: Homo sapiens
US-10-741-600-9856	US-10-741-600-9856	US-10-741-600-9891	US-10-741-600-9894
Query Match	78.2%; Score 17.2; DB 19; Length 201;	78.2%; Score 17.2; DB 19; Length 201;	78.2%; Score 17.2; DB 19; Length 201;
Best Local Similarity	86.4%; Pred. No. 3.8e+02;	86.4%; Pred. No. 3.8e+02;	86.4%; Pred. No. 3.8e+02;
Matches	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY	1 CACCTTCCTCCATCCTCAGAC 22	1 CACCTTCCTCCATCCTCAGAC 22	1 CACCTTCCTCCATCCTCAGAC 22
Db	71 CACCTTCCTCTCTCTCCTCAGAC 50	65 CACCTTCCTCTCTCTCCTCAGAC 44	89 CACCTTCCTCTCTCTCCTCAGAC 68
RESULT 38	US-10-741-600-9857/c	US-10-741-600-9891/c	

Qy 1 CACCTTCCTTCGATCCTCAGAC 22
| | | | | | | | | | | | | | | |
Db 139 CACCTTCCTTCCTCCTCAGAC 118

Search completed: June 4, 2005, 12:19:57
Job time : 210.716 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:32:00 ; Search time 1343.17 Seconds
(without alignments)
623.460 Million cell updates/sec

Title: US-09-674-277-22
Perfect score: 22
Sequence: 1 cacccttccttccttcagac 22

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20.4	92.7	329	9	CL385624 RPCI44_33
2	20	90.9	390	2	BE478771 163162 BA
3	20	90.9	416	1	AV668123 AV668123
4	20	90.9	430	1	AV596235 AV596235
5	20	90.9	837	7	CK774810 963604 MA
6	19	86.4	534	8	AZ341294 1M0073A11
7	19	86.4	726	9	AG40186 Mub muscu
8	19	86.4	1094	8	CC271799 CH261-130
9	18.8	85.5	398	1	AI006840 ua74f03.r
10	18.8	85.5	510	7	CF605975 RADIC01.0
11	18.8	85.5	530	8	BZ332151 bx25g04.g
12	18.8	85.5	535	8	BH771400 hg31g12.g
13	18.8	85.5	536	9	CG106750 PULK17TB
14	18.8	85.5	554	8	BZ789227 PUGBL32TD
15	18.8	85.5	555	2	BE899683 208237 MA
16	18.8	85.5	602	6	CB690142 CBST-51-F
17	18.8	85.5	641	1	AA946885 qc53d11.8
18	18.8	85.5	687	7	CF202696 R8890915N
19	18.8	85.5	705	7	CF213272 CGF100063
20	18.8	85.5	708	7	CF202739 R8890915N
21	18.8	85.5	729	7	CF213358 CGF100063
22	18.8	85.5	736	7	CF511363 CBud0002
23	18.8	85.5	741	9	CE272879 tigr-gss-
24	18.8	85.5	750	7	CF511446 CBud0002

C 25	18.8	85.5	761	7	CF203807
C 26	18.8	85.5	944	8	BZ789222
C 27	18.4	83.6	269	8	CC420222
C 28	18.4	83.6	367	1	AA607777
C 29	18.4	83.6	398	6	BY637300
C 30	18.4	83.6	398	6	BY637300
C 31	18.4	83.6	452	8	AZ062773
C 32	18.4	83.6	561	4	BZ555863
C 33	18.4	83.6	569	4	BZ554774
C 34	18.4	83.6	569	4	BZ564239
C 35	18.4	83.6	586	9	CL160038
C 36	18.4	83.6	667	7	CO099643
C 37	18.4	83.6	713	9	CL180445
C 38	18.4	83.6	733	7	CO146781
C 39	18.4	83.6	736	8	BZ936415
C 40	18.4	83.6	755	9	CL180446
C 41	18.4	83.6	778	9	CG832110
C 42	18.4	83.6	1134	8	CC244683
C 43	18	81.8	294	8	CC102721
C 44	18	81.8	578	8	BH235837
C 45	18	81.8	819	5	BX378582

ALIGNMENTS

RESULT 1
CL385624
LOCUS
DEFINITION
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CL385624
RPCI44_332M17.f RPCI-44 Sus scrofa genomic clone RPCI44_332M17,
genomic survey sequence.
CL385624
CL385624.1 GI:51437589
GSS.
Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Rogatcheva,M.B., Meyers,S., He,W., Larkin,D.M., Marron,B.M.,
Beever,J.E. and Schook,L.B.
Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
Through Comparative Genomics
Unpublished (2004)
Other_GSSs: RPCI44_332M17.r
Contact: Lawrence B. Schook
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 265 5326
Fax: 217 244 5617
Email: schook@uiuc.edu
Clones are derived from the porcine BAC library RPCI-44
(http://www.bacpac.chori.org/porcine242.htm). For BAC library
availability, please contact Pieter de Jong (pdjong@chori.org).
Clones may be purchased from BACPAC Resources
(http://BACPACResources.chori.org). This work was undertaken as part
of the International Swine Genome Sequencing Consortium by
University of Illinois at Urbana Champaign, USA with funds provided
by grant No. AG2002-34480-11828 from USDA-CSREES and
AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
Initiative)
Plate: 332 row: M column: 17
Seq primer: T7
Class: BAC ends.

FEATURES
source
1..329
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="four pigs (breed: 37.5% Yorks Landrace and 25%
Meishan)"
/db_xref="taxon:9823"
/clone="RPCI44_332M17"


```

JOURNAL      bovine ESTs
MEDLINE      Nucleic Acids Res. 29 (22), E108 (2001)
PUBMED       21570554
COMMENT      11713328
             Contact: Yoshikazu Sugimoto
             Animal Genetics Division
             Shirakawa Institute of Animal Genetics
             Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
             Tel: 81-248-25-5641
             Fax: 81-248-25-5725
             Email: kazuugi@cocoa.ocn.ne.jp
             Single pass sequencing.
             This clone was obtained from a polyA-deleted cDNA library.

FEATURES
    source
    Location/Qualifiers
        1..430
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            /clone="E1CA024B01"
            /tissue_type="cartilage"
            /dev_stage="fetus"
            /lab_host="DH10B"
            /notes="Vector: pZL1; Site 1: SalI; Site 2: NotI; Poly A
            was deleted from a NotI site"

ORIGIN
Query Match      90.9%; Score 20; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3  CCTTCCTTCATCCTCAGAC 22
        |||
Db      239  CCTTCCTTCATCCTCAGAC 220

RESULT 5
LOCUS      CK774810/c      837 bp      mRNA      linear      EST 20-FEB-2004
DEFINITION 963604 MARC 2B0V Bos taurus cDNA 3', mRNA sequence.
ACCESSION  CK774810
VERSION     CK774810.1 GI:42728953
KEYWORDS
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1 (bases 1 to 837)
AUTHORS    Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
            Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,
            Quackenbush,J. and Keale,J.W.
            Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
            Genome Res. 11 (4), 626-630 (2001)
JOURNAL     21180013
MEDLINE     11282978
PUBMED
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.ueda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: 105 row: C column: 4
            Seq primer: GTAATACGACTCACTATAGG.
FEATURES
    source
    Location/Qualifiers
        1..837
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            /lab_host="DH10B"
            /clone="E1CA024B01"
            /tissue_type="cartilage"
            /dev_stage="fetus"
            /lab_host="DH10B"
            /notes="Vector: pZL1; Site 1: SalI; Site 2: NotI; Poly A
            was deleted from a NotI site"

ORIGIN
Query Match      90.9%; Score 20; DB 7; Length 837;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3  CCTTCCTTCATCCTCAGAC 22
        |||
Db      255  CCTTCCTTCATCCTCAGAC 236

RESULT 6
AZ341294/c
LOCUS      AZ341294      534 bp      DNA      linear      GSS 29-SEP-2000
DEFINITION 1M0073A11R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0073A11 R, genomic survey sequence.
ACCESSION  AZ341294
VERSION     AZ341294.1 GI:10417402
KEYWORDS    GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
            Reilly,M., Rose,R., Stokes,R., Tingey,A., von
            Niederhausern,A. and Wright,D. Weiss,R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0073 row: A column: 11
            Seq primer: CACACAGGAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 534.
            Location/Qualifiers
                1..534
                    /organism="Mus musculus"
                    /mol_type="genomic DNA"
                    /strain="C57BL/6J"
                    /db_xref="taxon:10090"
                    /clone="UUGC1M0073A11"
                    /sex="Male"
                    /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                    /clone_lib="Mouse 10kb plasmid UUGC1M library"
                    /note="Vector: PWD42nv; Purified genomic DNA from M.
                    musculus C57BL/6J (male) was obtained from the Jackson
                    Laboratory Mouse DNA Resource
                    (http://www.jax.org/resources/documents/dnares/). The DNA
                    was hydrodynamically sheared by repeated passage through a
                    0.005 inch orifice at constant velocity. The sheared DNA
                    was blunt end-repaired with T4 DNA polymerase and T4
                    polynucleotide kinase. Adaptor oligonucleotides were
                    ligated to the blunt ends in high molar excess. The
                    adapted DNA was purified and size-selected for a 9.5 to

```

10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 86.4%; Score 19; DB 8; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCA 19
|||||
Db 353 CACCTTCCTTCCATCCTCA 335

RESULT 7

AG440186/c

LOCUS 726 bp DNA linear GSS 03-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-321C03.TJ, genomic survey sequence.

ACCESSION AG440186

VERSION AG440186.1 GI:48083249

KEYWORDS GSS.

SOURCE Mus musculus molossinus

ORGANISM Mus musculus molossinus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01

TITLE 2 (bases 1 to 726)

JOURNAL Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.

REFERENCE Direct Submission

TITLE Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou,Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:hattori@gsc.riken.jp, URL:http://hgsc.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the mouse BAC library MSMg01. For BAC

library availability, please contact Kuniya Abe (abe@rtc.riken.jp).

Teukuba Institute, Bio Resource Center,

The Institute of Physical and Chemical Research (RIKEN) 3-1-1

Koyadai, Tsukuba, 305-0074 Japan

phone: 81-298-36-9189, fax: 81-298-36-9199

e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : TJ

LIBRARY

Vector : pBACe3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

Location/Qualifiers

1. -726

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMg01-321C03.TJ"

/sex="male"

/tissue_type="mixture of kidney and spleen"

/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 86.4%; Score 19; DB 9; Length 726;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCCATCCTCA 19
|||||
Db 535 CACCTTCCTTCCATCCTCA 517

RESULT 8

CC271799

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE

AUTHORS

Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,

Warren,W., Graves,T., Mardis,E. and Wilson,R.

Gallus gallus BAC End Reads

Unpublished (2003)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Insert Length: 182000 Std Error: 0.00

Seq primer: Sp6 ATTAGGTGACACTATAG

Class: BAC ends

High quality sequence start: 29

High quality sequence stop: 564.

FEATURES

source

1..1094

Location/Qualifiers

/organism="Gallus gallus"

/mol_type="genomic DNA"

/strain="Red Jungle Fowl"

/db_xref="taxon:9031"

/clone="CH261-130E8"

/sex="female"

/cell_line="UCD001, inbred 256"

/clone_lib="CH261"

/note="Vector: PTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CH261 Female Chicken library - for library and clone

ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 86.4%; Score 19; DB 8; Length 1094;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCTTCCTTCCATCCTCAGA 21

|||||

Db 612 CCTTCCTTCCATCCTCAGA 630

RESULT 9

AI006840/c

LOCUS

DEFINITION

IMAGE:1363229 5', similar to gb:M93056 LEUKOCYTE ELASTASE INHIBITOR

(HUMAN); mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

CC271799 1094 bp DNA linear GSS 13-MAY-2003
CH261-130E8 Sp6.1 CH261 Gallus gallus genomic clone CH261-130E8,
genomic survey sequence.

ACCESSION CC271799

VERSION CC271799.1 GI:30626250

KEYWORDS GSS.

SOURCE Gallus gallus (chicken)

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE 1 (bases 1 to 1094)

AUTHORS Warren,W., Graves,T., Mardis,E. and Wilson,R.

Gallus gallus BAC End Reads

Unpublished (2003)

Contact: Richard K. Wilson

Genome Sequencing Center

Washington University School of Medicine

Email: submissions@watson.wustl.edu

Insert Length: 182000 Std Error: 0.00

Seq primer: Sp6 ATTAGGTGACACTATAG

Class: BAC ends

High quality sequence start: 29

High quality sequence stop: 564.

FEATURES

Location/Qualifiers

1..1094

/organism="Gallus gallus"

/mol_type="genomic DNA"

/strain="Red Jungle Fowl"

/db_xref="taxon:9031"

/clone="CH261-130E8"

/sex="female"

/cell_line="UCD001, inbred 256"

/clone_lib="CH261"

/note="Vector: PTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;

CH261 Female Chicken library - for library and clone

ordering information: http://www.chori.org/bacpac"

ORIGIN

Query Match 86.4%; Score 19; DB 8; Length 1094;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCTTCCTTCCATCCTCAGA 21

|||||

Db 612 CCTTCCTTCCATCCTCAGA 630

RESULT 9

AI006840/c

LOCUS

DEFINITION

IMAGE:1363229 5', similar to gb:M93056 LEUKOCYTE ELASTASE INHIBITOR

(HUMAN); mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)
 Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:896449
 Seq primer: -28ml3 rev2 ET from Amersham
 High quality sequence stop: 297.

FEATURES

source
 1. .398
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:1363229"
 /sex="male"
 /tissue_type="Thymus"
 /dev_stage="4 weeks"
 /lab_host="DH10B"

/clone_lib="Soares_thymus_2NBM7"
 /note="Vector: p773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTACCAACTGAGTGGAGCGCGCTTTTCTTTTCTTTTCTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified p773 vector. RNA
 provided by Dr. Bertrand Jordan. Library went through two
 rounds of normalization, and was constructed by Bento
 Soares and M.Patima Bonaldo."

ORIGIN

Query Match 85.5%; Score 18.8; DB 1; Length 398;
 Best Local Similarity 90.9%; Pred. No. 1.4e+03;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CACCTTCCTTCATCCTCAGAC 22
 |||||
 Db 366 CTCCTTCCTTCATCCTCAGAC 345

RESULT 10
 CF605975/c
 LOCUS
 DEFINITION
 RADIC01_001783 Grape Root pSPORT1 Library Vitis vinifera cDNA 5',
 mRNA sequence.

ACCESSION
 VERSION
 CF605975.1 GI:37186622
 EST.
 KEYWORDS
 SOURCE
 ORGANISM

Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.

REFERENCE
 AUTHORS
 Moser, C., Segala, C., Fontana, P., Salakhudinov, I., Gatto, P.,
 Pindo, M., Zyprian, E., Toepfer, R., Grando, M.S. and Velasco, R.
 Titled
 Expressed sequence tags from different organs of Vitis vinifera
 Unpublished (2003)
 Contact: Moser C

Laboratorio di Genetica Molecolare
 Istituto Agrario di San Michele all'Adige (IASMA)
 via E. Mach 1, San Michele all'Adige (TN), I-38010, Italia
 Tel: 0039-0461-615314
 Fax: 0039-0461-650956

Email: claudio.moser@ismaa.it
 The sequencing work has been funded by the 'Fondazione Cassa di
 Risparmio di Trento e Rovereto'
 High quality sequence stop: 510.

FEATURES

source
 1. .510
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Pinot Noir"
 /db_xref="taxon:29760"
 /sex="Hermaphrodite"
 /dev_stage="young root (first year)"
 /lab_host="DH10B"
 /clone_lib="Grape Root pSPORT1 Library"
 /note="Organ: root; Vector: pSPORT1; Site_1: NotI; Site_2:
 Sall"

ORIGIN

Query Match 85.5%; Score 18.8; DB 7; Length 510;
 Best Local Similarity 90.9%; Pred. No. 1.4e+03;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 CACCTTCCTTCATCCTCAGAC 22
 |||||
 Db 507 CACTTCCTTCATCCTCAGAC 486

RESULT 11

BZ332151
 LOCUS
 DEFINITION
 hx25g04.g1 WGS-Sbicolorp (JM107 adapted methyl filtered) Sorghum
 bicolor genomic clone hx25g04 5', genomic survey sequence.

ACCESSION
 VERSION
 BZ332151
 GSS.
 KEYWORDS
 SOURCE
 ORGANISM

Sorghum bicolor (sorghum)
 Sorghum bicolor
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
 clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE

AUTHORS
 Rabinowicz, P.D., O'Shaughnessy, A.L., Balija, V., Dedhia, N.,
 Katzenburger, P., King, L., Miller, B., Muller, S., Nascimento, L.,
 Zutavern, T., Palmer, L., McCombie, W.R. and Martienssen, R.A.
 Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
 Unpublished (2002)
 Contact: W. Richard McCombie

TITLE

JOURNAL
 COMMENT
 Lita Annenberg Hazen Genome Sequencing Center
 Cold Spring Harbor Laboratory
 PO Box 100, Cold Spring Harbor, NY 11724, USA
 Tel: 516 367 8884
 Fax: 516 367 8874
 Email: mcombie@cshl.org
 Plate: hx25 row: 9 column: 04
 Seq primer: -21M13UnivRev
 Class: shotgun
 High quality sequence stop: 530.

FEATURES

source
 1. .530
 /organism="Sorghum bicolor"
 /mol_type="genomic DNA"
 /db_xref="taxon:4558"
 /clone="hx25g04"
 /lab_host="JM107 or DH5a"
 /clone_lib="WGS-Sbicolorp (JM107 adapted methyl filtered)"
 /note="Site 1: Xba I; Site 2: Xba I; The vector was
 digested with XbaI and one nucleotide was added by fill in
 in the recessive 3' end. The genomic DNA was nebulized,
 end repaired, adaptor ligated and size fractionated using
 sephadex. The resulting fragments were between 0.8 and 3
 kb and were cloned into the vector (.x/y reads in M13mp19,
 .b/g reads in pUC19). The same ligation was transformed in
 either JM107 or DH5a."

ORIGIN

Query Match 85.5%; Score 18.8; DB 8; Length 530;
 Best Local Similarity 90.9%; Pred. No. 1.4e+03;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCAGAC 22
 |||||
 Db 341 CACCTTCCTTCATCTCAGAC 362

RESULT 12

BH771400

LOCUS

DEFINITION hg31g12-g1 JM107 adapted methyl filtered library Zea mays genomic

clone hg31g12 5', genomic survey sequence.

BH771400

ACCESSION

VERSION

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 535)

Rabinowicz,P.D., Schutz,K., Dedhia,N., Yordan,C., Parnell,L.D.,

Stein,L., McCombie,W.R. and Martienssen,R.A.

Differential methylation of genes and retrotransposons allows

shotgun sequencing of the maize genome

Nat. Genet. 23, 305-308 (1999)

Contact: W. Richard McCombie

Lita Annenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: hg31 row: g column: 12

Seq primer: M13 Universal reverse sequencing primer

Class: shotgun

High quality sequence stop: 535.

Location/Qualifiers

1..535

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="hg31g12"

/tissue_type="Tissue type"

/lab_host="JM107"

/clone_lib="JM107 adapted methyl filtered library"

I; Site 2: Xba I; DNA prepared from purified nuclei was

randomly sheared by nebulization, size fractionated to

enrich for the 0.5 to 4 kbp fraction, end repaired and

ligated using compatible adaptors into Xba I digested M13

vector filled-in to produce a 3 bp protruding 5' end. The

ligation mix was electroporated into E.coli JM107MA2."

ORIGIN

Query Match 85.5%; Score 18.8; DB 8; Length 535;

Best Local Similarity 90.9%; Pred. No. 1.4e+03;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCAGAC 22

|||||

Db 454 CACCTTCCTTCAGCCCCAGAC 475

RESULT 13

CG106750/c

LOCUS

DEFINITION FUILLK17TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0610C09,

536 bp DNA linear GSS 20-AUG-2003

CG106750 536 bp DNA linear GSS 20-AUG-2003

FUILLK17TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0610C09,

genomic survey sequence.

CG106750

VERSION

KEYWORDS

SOURCE

ORGANISM

Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 536)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and

Bennetzen,J.

Maize Genomics Consortium

Unpublished (2003)

Other_GSSs: PULLK17TD

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..536

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBTa0610C09"

/clone_lib="ZM_0.6_1.0_KB"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high

CoT selected genomic DNA library"

ORIGIN

Query Match 85.5%; Score 18.8; DB 9; Length 536;

Best Local Similarity 90.9%; Pred. No. 1.4e+03;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCTCAGAC 22

|||||

Db 203 CTCCTTCCTTCATCTCAGAC 182

RESULT 14

BZ789227

LOCUS

DEFINITION

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

BZ789227

```

FEATURES
  source      Location/Qualifiers
1..554
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone_lib="ZM0.6_1.0.KB"
  /notes="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
  Cor selected genomic DNA library"

ORIGIN
Query Match      85.5%; Score 18.8; DB 8; Length 554;
Best Local Similarity 90.9%; Pred. No. 1.4e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
    |||||
Db 351 CTCCTTCCTTCATCCTCAGAC 372

RESULT 15
LOCUS      BE899683
DEFINITION 208237 MARC 2BOV Bos taurus CDNA 5', mRNA sequence.
ACCESSION  BE899683
VERSION     BE899683.1 GI:10387255
KEYWORDS   EST.
SOURCE      Bos taurus (cow)
ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
REFERENCE   1 (bases 1 to 555)
  AUTHORS   Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
            Bennett,G.L., Heaton,M.P., Laegreid,W., Rohrer,G.A.,
            Chitko-McKown,C.G., Fertea,G., Holt,I., Karamycheva,S., Liang,P.,
            Quackenbush,J. and Keefe,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   1180013
  PUBMED    11282978
  COMMENT   Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -mismatch 12 options.
            PCR Primers
            FORWARD: AGGACACGCTATGACCAT
            BACKWARD: GTTTCACGTCAGCAGC
            Plate: 55 row: K column: 3
            Seq primer: ATTAGTGACACTAG.
FEATURES
  source      Location/Qualifiers
1..555
  /organism="Bos taurus"
  /mol_type="mRNA"
  /db_xref="taxon:9913"
  /tissue_type="pooled"
  /lab_host="DH10B"
  /clone_lib="MARC 2BOV"
  /notes="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
  Library made from pooled tissue from testis, thymus,
  semitendinosus muscle, longissimus muscle, pancreas,
  adrenal, and endometrium."

ORIGIN
Query Match      85.5%; Score 18.8; DB 2; Length 555;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
    |||||
Db 395 CACCTTCCTTCATCCTCAGAC 374

RESULT 17
LOCUS      AA946885/c
DEFINITION 208237 MARC 2BOV Bos taurus CDNA 5', mRNA sequence.
ACCESSION  AA946885
VERSION     AA946885.1 GI:3110280
  
```

```

Best Local Similarity 90.9%; Pred. No. 1.4e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
    |||||
Db 182 CCCCTTCCTTCCTCTCAGAC 203

RESULT 16
LOCUS      CB690142/c
DEFINITION 602 bp mRNA linear EST 04-SBP-2003
            CEST-51-F-07 Mixed source, strain EP155 and EP155 infected with
            hypovirus CHV1-EP713 Cryphonectria parasitica cDNA clone EP155,
            EP155-CHV1-EP713 5-prime, mRNA sequence.
ACCESSION  CB690142
VERSION     CB690142.1 GI:34451011
KEYWORDS   EST.
SOURCE      Cryphonectria parasitica
            Cryphonectria parasitica
            Cryphonectria parasitica
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Sordariomycetidae; Diaporthales; Valsaceae; Cryphonectria-Endothia
            complex; Cryphonectria.
REFERENCE   1 (bases 1 to 602)
  AUTHORS   Dawe,A.L., McMains,V.C., Panglao,M., Kasahara,S., Chen,B. and
            Nuss,D.L.
  TITLE     An ordered collection of expressed sequences from Cryphonectria
            parasitica and evidence of genomic microsynteny with Neurospora
            crassa and Magnaporthe grisea
  JOURNAL   Microbiology 149 (9), 2373-2384 (2003)
  MEDLINE   22830414
  PUBMED    12949163
  COMMENT   Contact: Dawe AL, Nuss DL
            Center for Biosystems Research
            University of Maryland Biotechnology Institute
            5115 Plant Sciences Building, College Park, MD 20742, USA
            Tel: 301-405-7661 or 5111
            Fax: 301-314-9075
            Email: dawe@cumbi.umd.edu
            Cryphonectria parasitica EST derived from mixed samples of cDNA
            from wild-type (strain EP155) and hypovirus-infected cultures.
            Seq primer: M13-reverse.
FEATURES
  source      Location/Qualifiers
1..602
  /organism="Cryphonectria parasitica"
  /mol_type="mRNA"
  /strain="EP155"
  /db_xref="ATCC (inhost):38755 and 52571"
  /db_xref="taxon:5116"
  /clone="EP155, EP155-CHV1-EP713"
  /clone_lib="Mixed source, strain EP155 and EP155 infected
  with hypovirus CHV1-EP713"
  /note="Cryphonectria parasitica EST derived from mixed
  samples of cDNA from wild-type (strain EP155) and
  hypovirus-infected cultures"

ORIGIN
Query Match      85.5%; Score 18.8; DB 6; Length 602;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
    |||||
Db 395 CACCTTCCTTCATCCTCAGAC 374

RESULT 17
LOCUS      AA946885/c
DEFINITION 641 bp mRNA linear EST 23-JUL-1998
            oq53d11.s1 NCI_CGAP Kids Homo sapiens cDNA clone IMAGE:1590069 3'
            similar to gb:J02947 EXTRACELLULAR SUPEROXIDE DISMUTASE PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION  AA946885
VERSION     AA946885.1 GI:3110280
  
```

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 641)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-@email.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 707 Std Error: 0.00
Seq primer: -40ml3 fwd. RT from Amersham
High quality sequence stop: 331.
FEATURES
source
Location/Qualifiers
1..641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1590069"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid5"
/note="Organ: Kidney; Vector: p7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
ACTGAGAGATTCGGCGGCAGATATTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "
ORIGIN
Query Match 85.5%; Score 18.8; DB 1; Length 641;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CACCTTCCTCCATCCTCAGAC 22
Db 431 CACCATCCTCCATCCTCAGAC 410
RESULT 18
LOCUS CF202696 687 bp mRNA linear EST 04-AUG-2003
DEFINITION RR890915N0006_I1a_Fa_G06 Vitis sp. RR890915N Vitis hybrid cultivar
cDNA clone RR890915N0006_I1a_Fa_G06 5', mRNA sequence.
ACCESSION CF202696
VERSION CF202696.1 GI:33397069
KEYWORDS EST.
SOURCE Vitis hybrid cultivar
ORGANISM Vitis hybrid cultivar
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 687)
AUTHORS Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Leslie,A., Xu,J.,
Jones,K. and Cook,D.
TITLE Transcriptional responses of a Pierce's Disease resistant Vitis sp.
to infection by Xylella fastidiosa
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: GCCAACGATGGTCTAG.
Location/Qualifiers
1..705
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAST0003_I1R_C12"
/sex="Hermaphrodite"
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: ACGTACCGACATATGCC.
Location/Qualifiers
1..687
/organism="Vitis hybrid cultivar"
/mol_type="mRNA"
/db_xref="taxon:241073"
/clone="RR890915N0006_I1a_Fa_G06"
/lab_host="DH5alpha"
/clone_lib="Vitis sp. RR890915N"
/note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:
SfiI; RR890915N is a cDNA library of leaves from the Fl of
Vitis rupestris 'A. de Serres' x V. spp. 'B42-26',
(8909-15) showing resistance to Xylella fastidiosa under
greenhouse experimental conditions (M. Andrew Walker and
Alan Krivanek, U C Davis). Samples were collected from
17-week old greenhouse grown plants. cDNAs were made by
oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAACGAGTGGCCATTACGCGCGG-3' and
5'-ATTCTAGAGCGCGCGGCGGACATG-dT(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."
FEATURES
source
Location/Qualifiers
1..687
/organism="Vitis hybrid cultivar"
/mol_type="mRNA"
/db_xref="taxon:241073"
/clone="RR890915N0006_I1a_Fa_G06"
/lab_host="DH5alpha"
/clone_lib="Vitis sp. RR890915N"
/note="Organ: Leaf; Vector: pDNR; Site 1: SfiI; Site 2:
SfiI; RR890915N is a cDNA library of leaves from the Fl of
Vitis rupestris 'A. de Serres' x V. spp. 'B42-26',
(8909-15) showing resistance to Xylella fastidiosa under
greenhouse experimental conditions (M. Andrew Walker and
Alan Krivanek, U C Davis). Samples were collected from
17-week old greenhouse grown plants. cDNAs were made by
oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAACGAGTGGCCATTACGCGCGG-3' and
5'-ATTCTAGAGCGCGCGGCGGACATG-dT(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."
ORIGIN
Query Match 85.5%; Score 18.8; DB 7; Length 687;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CACCTTCCTCCATCCTCAGAC 22
Db 505 CACTTTCCTCCATCCTCCGAC 484
RESULT 19
LOCUS CF213272 705 bp mRNA linear EST 01-AUG-2003
DEFINITION CGF1000632_C12 Vitis vinifera cv. cabernet sauvignon Stem - CAST
Vitis vinifera cDNA clone CAST0003_I1R_C12 3', mRNA sequence.
ACCESSION CF213272
VERSION CF213272.1 GI:33407645
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 705)
AUTHORS Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Leslie,A., Xu,J.,
Jones,K. and Cook,D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: GCCAACGATGGTCTAG.
Location/Qualifiers
1..705
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAST0003_I1R_C12"
/sex="Hermaphrodite"

```



```

RESULT 22
CF511363/c
LOCUS
DEFINITION
  CF511363 736 bp mRNA linear EST 09-SEP-2003
  Cabud0002_IIF_F04 Vitis vinifera cv. cabernet sauvignon (Clone 8)
  Bud - CABUD Vitis vinifera cDNA clone Cabud0002_IIF_F04 5', mRNA
  sequence.
ACCESSION
CF511363.1 GI:34543131
VERSION
KEYWORDS
SOURCE
ORGANISM
  Vitis vinifera
  Vitis vinifera

REFERENCE
AUTHORS
  Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Leslie,A., Xu,J.,
  Jones,K. and Cook,D.
TITLE
  Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
  berries at various developmental stages
JOURNAL
COMMENT
  Contact: Douglas Cook, PhD
  CAES Genome Facility
  UC Davis, Plant Pathology
  One Shields Ave, Davis, CA 95616, USA
  Tel: 530 754 6561
  Fax: 530 754 6617
  Email: drcook@ucdavis.edu
  Seq primer: ACGTACCGGACATATGCC.
  Location/Qualifiers
  1..736
  /organism="Vitis vinifera"
  /mol_type="mRNA"
  /cultivar="Cabernet Sauvignon (Clone 8)"
  /db_xref="taxon:29760"
  /clone="CABUD0002_IIF_F04"
  /sex="Hermaphrodite"
  /dev_stage="Pre-bloom (10-11 days before bloom)"
  /lab_host="DH5alpha"
  /clone_lib="Vitis vinifera cv. cabernet sauvignon (Clone
  8) Bud - CABUD"
  /note="Organ: Bud; Vector: pDNR; Site_1: SfiI; Site_2:
  SfiI; CABUD is a cDNA library of Vitis vinifera cv.
  'Cabernet Sauvignon' clone 8 dissected buds. Samples were
  collected May 13, 2002 from pre-bloom plants (10-11 days
  before bloom), pre-veraison. Sampled vines were located at
  the University of California, Davis, Experimental
  Vineyard. cDNAs were made by oligo-dT priming and
  directionally cloned. 5' and 3' adaptors were used in
  cloning as follows:
  5'-AAGCAGTGTATCAACGAGAGTGCCATTACGCGGG-3' and
  5'-ATTCTAGGCGCGAGCGCGCATG-dT(30)NN-3'. Library was
  constructed using the Clontech Creator SMART kit and
  size-selected to contain the 0.5-3 kb size fraction."

ORIGIN
Query Match 85.5%; Score 18.8; DB 7; Length 736;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
|||||
Db 511 CACTTCCTTCATCCTCAGAC 490
|||||

RESULT 23
CE272879/c
LOCUS
DEFINITION
  CE272879 741 bp DNA linear GSS 26-SEP-2003
  tigr-gss-dog-17000333573565 Dog Library Canis familiaris genomic,
  genomic survey sequence.
ACCESSION
CE272879.1 GI:36008717
VERSION
KEYWORDS
SOURCE
ORGANISM
  Canis familiaris (dog)
  Canis familiaris
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
  1 (bases 1 to 741)
  Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
  Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
  Venter,J.C.
  The dog genome: survey sequencing and comparative analysis
  Science 301 (5641), 1898-1903 (2003)
  22875432
  14512627
  COMMENT
  Contact: Kirkness EF
  The Institute for Genomic Research
  Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
  Rockville, MD 20850, USA
  Tel: 301-838-0200
  Fax: 301-838-0208
  Email: ekirknes@tigr.org
  Class: shotgun.
  Location/Qualifiers
  1..741
  /organism="Canis familiaris"
  /mol_type="genomic DNA"
  /strain="Standard Poodle"
  /db_xref="taxon:9615"
  /clone_lib="Dog Library"
  /note="Site 1: BstXI; Libraries were prepared from
  peripheral blood"

ORIGIN
Query Match 85.5%; Score 18.8; DB 9; Length 741;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CACCTTCCTTCATCCTCAGAC 22
|||||
Db 262 CACTTCCTTCATCCTCTGTC 241
|||||

RESULT 24
CF511446
LOCUS
DEFINITION
  CF511446 750 bp mRNA linear EST 09-SEP-2003
  Cabud0002_IIR_F04 Vitis vinifera cv. cabernet sauvignon (Clone 8)
  Bud - CABUD Vitis vinifera cDNA clone Cabud0002_IIR_F04 3', mRNA
  sequence.
ACCESSION
CF511446.1 GI:34543214
VERSION
KEYWORDS
SOURCE
ORGANISM
  Vitis vinifera
  Vitis vinifera
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; Vitaceae; Vitis.
  1 (bases 1 to 750)
  Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Leslie,A., Xu,J.,
  Jones,K. and Cook,D.
  Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
  berries at various developmental stages
  Unpublished (2003)
  Contact: Douglas Cook, PhD
  CAES Genome Facility
  UC Davis, Plant Pathology
  One Shields Ave, Davis, CA 95616, USA
  Tel: 530 754 6561
  Fax: 530 754 6617
  Email: drcook@ucdavis.edu
  Seq primer: GCCAACGATGCTCTAG.
  Location/Qualifiers
  1..750
  /organism="Vitis vinifera"
  /mol_type="mRNA"

```

```

/cultivar="Cabernet Sauvignon (Clone 8)"
/db_xref="taxon:29760"
/clone="Cabud0002_IIR_F04"
/sex="Hermaprodite"
/dev_stage="Pre-bloom (10-11 days before bloom)"
/lab_host="DH5alpha"
/clone_l1b="Vitis vinifera cv. cabernet sauvignon (Clone
8) Bud - CABUD"
/notes="Organ: Bud; Vector: pDNR; Site_1: Sfil; Site_2:
Sfil; CABUD is a cDNA library of Vitis vinifera cv.
'Cabernet Sauvignon' Clone 8 dissected buds. Samples were
collected May 13, 2002 from pre-bloom plants (10-11 days
before bloom), pre-veraison. Sampled vines were located at
the University of California, Davis, Experimental
vineyard. cDNAs were made by oligo-dT priming and
directionally cloned. 5' and 3' adaptors were used in
cloning as follows:
5'-AAGCATGTGTATCAGCGAGTGGCCATGACGGCGGG-3' and
5'-ATTCTAGAGCGAGCGCCAGCATG-DT(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

```

```

ORIGIN
size-selected to contain the 0.3-5 kb size fraction.

Query Match      85.5%; Score 18.8; DB 7; Length 750;
Best Local Similarity 90.9%; Pred.No.1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 CACCTTCCTTCCATCCTCTCGAGAC 22
    ||| ||||| ||||| |||||
Db   226 CACTTTGCTTCCATCCTCTCGAGAC 247

```

RESULT 25	CF203807/c	CF203807	761 bp	mRNA	linear	EST 04-AUG-2003
LOCUS						
DEFINITION		RR89091510001	Iva Fa G03	Vitis sp.	RR8909151	Vitis hybrid cultivar
		cDNA clone	RR89091510001	Iva Fa G03 5'	mRNA sequence.	

ACCESSION	CF203807
VERSION	CF203807.1
KEYWORDS	GI:33398180
SOURCE	EST.
ORGANISM	Vitis hybrid cultivar Vitis hybrid cultivar Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis. 1 (bases 1 to 761)
REFERENCE	Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Leslie,A., Xu,J., Jones,K. and Cook,D.
AUTHORS	Transcriptional responses of a Pierce's Disease resistant Vitis sp. to infection by Xylella fastidiosa
TITLE	Unpublished (2003)
JOURNAL	

UNPUBLISHED (2003)
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave. Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drccook@ucdavis.edu
Seq primer: ACGGTACGGACATATGCC.

```

seq primer: AccorAccessionInfo.
Location/Qualifiers
1. .761
/organism="Vitis hybrid cultivar"
/mol_type="mRNA"
/db_xref="taxon:241073"
/clones="RR89091510001_IVa_Fa_G03"
/lab_host="DH5alpha"
/clone_lib="Vitis sp. RR8909151"
/note="Organ: Leaf; Vector: pDNR; Site_1: Sfil; Site_2: Sfil; RR8909151 is a cDNA library of leaves from the F1 of Vitis rupestris 'A. de Serres' x V. spp. 'b42-26' (8909-15) showing resistance to Xylella fastidiosa under greenhouse experimental conditions (M. Andrew Walker and Alan Krivanek, U C Davis). Samples were collected from

```

17-week old greenhouse grown plants, nine weeks after inoculation with the bacterial pathogen *Xylella fastidiosa*. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:

5'-AAGCAGATGTCTACACGACATGCGCCATTACGGCGGG-3' and 5'-ATTCTAGAGCGGCGGCGGCAGATG-T(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

```

Query Match      85.5%; Score 18.8; DB 7; Length 761;
Best Local Similarity 90.9%; Pred. No. 1.5e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CACCTTCCTTCCTCCATCCTCCAGAC 22
Db 491 CACTTTTCCTTCCTCCATCCTCCGAC 470

```

RESULT 26
BZ789222/c

LOCUS	BZ789222	944 bp	DNA	linear	GSS 17-MAR-2003
DEFINITION	PUGBL32TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBta330E16, genomic survey sequence.				
ACCESSION	BZ789222				
VERSION	BZ789222.1	GI:28983476			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				

REFERENCE AUTHORS	TITLE JOURNAL COMMENT
1. J. H.
2. J. H.
3. J. H.
4. J. H.
5. J. H.
6. J. H.
7. J. H.
8. J. H.
9. J. H.
10. J. H.
11. J. H.
12. J. H.
13. J. H.
14. J. H.
15. J. H.
16. J. H.
17. J. H.
18. J. H.
19. J. H.
20. J. H.
21. J. H.
22. J. H.
23. J. H.
24. J. H.
25. J. H.
26. J. H.
27. J. H.
28. J. H.
29. J. H.
30. J. H.
31. J. H.
32. J. H.
33. J. H.
34. J. H.
35. J. H.
36. J. H.
37. J. H.
38. J. H.
39. J. H.
40. J. H.
41. J. H.
42. J. H.
43. J. H.
44. J. H.
45. J. H.
46. J. H.
47. J. H.
48. J. H.
49. J. H.
50. J. H.
51. J. H.
52. J. H.
53. J. H.
54. J. H.
55. J. H.
56. J. H.
57. J. H.
58. J. H.
59. J. H.
60. J. H.
61. J. H.
62. J. H.
63. J. H.
64. J. H.
65. J. H.
66. J. H.
67. J. H.
68. J. H.
69. J. H.
70. J. H.
71. J. H.
72. J. H.
73. J. H.
74. J. H.
75. J. H.
76. J. H.
77. J. H.
78. J. H.
79. J. H.
80. J. H.
81. J. H.
82. J. H.
83. J. H.
84. J. H.
85. J. H.
86. J. H.
87. J. H.
88. J. H.
89. J. H.
90. J. H.
91. J. H.
92. J. H.
93. J. H.
94. J. H.
95. J. H.
96. J. H.
97. J. H.
98. J. H.
99. J. H.
100. J. H.

TABLE 10.10

FEATURES

```

/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTA30E16"
/notes="Vector: pCR3-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

```

ORIGIN

```

Query Match      85.5%; Score 18.8; DB 8; Length 944;
Best Local Similarity 90.9%; Pred. No. 1.6e+03;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CACCTTCCTTCCTCCATCCTCAGAC 22
Db . 307 CTGCTTCCTTCCTCCATCCTCAGAC 286

```

RESULT 27
CC420222
LOCUS
DEFINITION

	CC420222	PUEG45TD	ZM 0.6	ZM 1.0	ZM 1.0 KB	Zea may's	genomic clone	ZMMBTa215H17,	DNA	linear	GSS 19-MAY-2003
									269 bp		

```

ACCESSION      CC420222
VERSION        CC420222.1  GI:30900312
KEYWORDS       GSS.
SOURCE         Zea mays
ORGANISM       Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 269)
AUTHORS        Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
               Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
               Bennetzen,J.
TITLE          Maize Genomics Consortium
JOURNAL        Unpublished (2003)
COMMENT        Contact: Cathy Whitelaw
               TIGR
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301-838-5943
               Fax: 301-838-0208
               Email: whitelaw@tigr.org
               Seq primer: TP
               Class: sheared ends.
FEATURES       Location/Qualifiers
               1..269
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone="ZMMBT245H17"
               /clone_lib="ZM_0.6_1.0_KB"
               /notes="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
               CoT selected genomic DNA library"
ORIGIN
Query Match      83.6%; Score 18.4; DB 8; Length 269;
Best Local Similarity 95.0%; Pred. No. 1.9e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  CCTTCTTCATCTCCTCAGAC 22
        |||||
Db      111 CCTTCTTCATCTCCTTAGAC 130

RESULT 28
AA607777
LOCUS          AA607777              367 bp      mRNA      linear      EST 30-SEP-1997
DEFINITION    Y08g12.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:1040902 5', mRNA sequence.
ACCESSION     AA607777
VERSION       AA607777.1  GI:2455212
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 367)
AUTHORS       Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
               Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
               Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
               Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
               Waterston,R.
TITLE         The WashU-HMI Mouse EST Project
JOURNAL       Unpublished (1996)
COMMENT       Contact: Marra M/Mouse EST Project
               WashU-HMI Mouse EST Project
               Washington University School of MedicineP
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: mouseest@watson.wustl.edu
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.

Genomic survey sequence.
CC420222
VERSION        CC420222.1  GI:30900312
KEYWORDS       GSS.
SOURCE         Zea mays
ORGANISM       Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE      1 (bases 1 to 269)
AUTHORS        Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
               Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
               Bennetzen,J.
TITLE          Maize Genomics Consortium
JOURNAL        Unpublished (2003)
COMMENT        Contact: Cathy Whitelaw
               TIGR
               9712 Medical Center Drive, Rockville, MD 20850, USA
               Tel: 301-838-5943
               Fax: 301-838-0208
               Email: whitelaw@tigr.org
               Seq primer: TP
               Class: sheared ends.
FEATURES       Location/Qualifiers
               1..269
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone="ZMMBT245H17"
               /clone_lib="ZM_0.6_1.0_KB"
               /notes="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
               CoT selected genomic DNA library"
ORIGIN
Query Match      83.6%; Score 18.4; DB 8; Length 269;
Best Local Similarity 95.0%; Pred. No. 1.9e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  CCTTCTTCATCTCCTCAGAC 22
        |||||
Db      111 CCTTCTTCATCTCCTTAGAC 130

RESULT 28
AA607777
LOCUS          AA607777              367 bp      mRNA      linear      EST 30-SEP-1997
DEFINITION    Y08g12.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:1040902 5', mRNA sequence.
ACCESSION     AA607777
VERSION       AA607777.1  GI:2455212
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 367)
AUTHORS       Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
               Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
               Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
               Theising,B., Wyllie,T., Lennon,G., Soares,B., Wilson,R. and
               Waterston,R.
TITLE         The WashU-HMI Mouse EST Project
JOURNAL       Unpublished (1996)
COMMENT       Contact: Marra M/Mouse EST Project
               WashU-HMI Mouse EST Project
               Washington University School of MedicineP
               4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
               Tel: 314 286 1800
               Fax: 314 286 1810
               Email: mouseest@watson.wustl.edu
               This clone is available royalty-free through LLNL ; contact the
               IMAGE Consortium (info@image.llnl.gov) for further information.

MG1:580926
Seq primer: -28mi3 rev1 ET from Amersham
High quality sequence stop: 362.
Location/Qualifiers
1..367
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:1040902"
/sex="females"
/tissue_type="whole skin"
/dev_stage="11 weeks old"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene mouse skin (#937313)"
/note="Organ: skin; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. Whole skin from 11 week old C57BL/6 female mice.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
ORIGIN
Query Match      83.6%; Score 18.4; DB 1; Length 367;
Best Local Similarity 95.0%; Pred. No. 2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  CCTTCTTCATCTCCTCAGAC 22
        |||||
Db      231 CCTTCTTCATCTCCTCAGAC 250

RESULT 29
BY637300
LOCUS          BY637300              398 bp      mRNA      linear      EST 15-DEC-2002
DEFINITION    BY637300 RIKEN full-length enriched, visual cortex Mus musculus
CDNA clone K430312D13 3', mRNA sequence.
ACCESSION     BY637300
VERSION       BY637300.1  GI:26972482
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 398)
AUTHORS       Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
               Nikaido,I., Oato,N., Saito,R., Suzuki,H., Yamanaka,I.,
               Kiyosawa,H., Yagi,K., Iomaru,Y., Hasegawa,Y., Nogami,A.,
               Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,
               Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,
               Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Bruscia,V.,
               Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,
               Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,
               Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,
               Kusajich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,
               Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,
               Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,
               Maltchis,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,
               Numata,K., Okido,T., Pavan,W.J., Perteza,G., Pesole,G.,
               Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,
               Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,
               Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shinada,K.,
               Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,
               Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y.,
               Wells,C., Wilming,I.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,
               Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,
               Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,
               Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,
               Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,
               Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,
               Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,
               Rogers,J., Birney,E. and Hayashizaki,Y.
               Analysis of the mouse transcriptome based on functional annotation

```


JOURNAL MEDLINE PUBMED COMMENT	of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 22354683 12466851		AUTHORS	
	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Michela Pagliolini and Takao K. Hensch (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hiroawa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge. Please visit our web site (http://genome.gsc.riken.go.jp) for further details.		TITLE	
	Query Match 83.6%; Score 18.4; DB 6; Length 398; Best Local Similarity 95.0%; Pred. No. 2e+03; Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		JOURNAL MEDLINE PUBMED COMMENT	
	2 ACCTCTCTCCATCCTCAGA 21 34 ACCTTCATCCATCCTCAGA 53			
RESULT 30 BY637346 LOCUS DEFINITION BY637346 RIKEN full-length enriched, visual cortex Mus musculus CDNA clone K430312115 3', mRNA sequence. ACCESSION BY637346 VERSION BY637346.1 GI:26972528 KEYWORDS EST. Mus musculus (house mouse) Mus musculus ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 398)	398 bp mRNA linear EST 15-DEC-2002 visual cortex Mus musculus		FEATURES source Location/Qualifiers 1. .398 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="K430312p13" /tissue type="visual cortex" /clone_lib="RIKEN full-length enriched, visual cortex"	
	2 ACCTCTCTCCATCCTCAGA 21			
	34 ACCTTCATCCATCCTCAGA 53			

```

/db_xref="taxon:10090"
/clone="R430312115"
/tissue type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"

```

ORIGIN

```

Query Match      83.6%; Score 18.4; DB 6; Length 398;
Best Local Similarity 95.0%; Pred. No. 2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  2 ACCTTCCTCCATCCTCAGA 21
    |||||
Db   33 ACCTTCCTCCATCCTCAGA 52

```

```

RESULT 31
AZ062773
LOCUS
DEFINITION
  AZ062773 452 bp DNA linear GSS 30-MAR-2000
  RPCI-23-415P2-TV RPCI-23 Mus musculus genomic clone RPCI-23-415P2,
  genomic survey sequence.
ACCESSION
  AZ062773
VERSION
  AZ062773.1 GI:7354022
KEYWORDS
  GSS.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 452)
REFERENCE
  Zhao,S., Niernan,W., Feldblum,T., Malek,J., Shatsman,S., de
  Akinret,B., Levine,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de
  Jong,P. and Fraser,C.M.
  Mouse BAC End Sequences from Library RPCI-23
  Unpublished (1999)
Other GSSs: RPCI-23-415P2.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

```

```

Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html
Plate: 415 row: P column: 2
Seq primer: T7
Class: BAC ends.

```

FEATURES

```

source
  Location/Qualifiers
    1..452
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="RPCI-23-415P2"
      /sex="Female"
      /lab_host="DH10B"
      /clone_lib="RPCI-23"
      /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
      EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
      brain genomic DNA was isolated and partially digested
      with a combination of EcoRI and EcoRI Methylase. Size
      selected DNA was cloned into the pBACe3.6 vector at the
      EcoRI sites. The ligation products were transformed into
      DH10B electrocompetent cells (BRL Life Technologies)."
```

ORIGIN

```

Query Match      83.6%; Score 18.4; DB 8; Length 452;
Best Local Similarity 95.0%; Pred. No. 2.1e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY

```

2 ACCTTCCTTCATCCTCAGA 21
  |||||
Db  409 ACGTTCTTCATCCTCAGA 428

```

RESULT 32

```

BJ555863/c
LOCUS
DEFINITION
  BJ555863
  CDNA clone jmi8h18 5', mRNA sequence.
ACCESSION
  BJ555863
VERSION
  BJ555863.1 GI:27237683
KEYWORDS
  EST.
SOURCE
  Ipomoea nil (Japanese morning glory)
ORGANISM
  Ipomoea nil
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Convolvulaceae; Ipomoeae; Ipomoea.
  1 (bases 1 to 561)
REFERENCE
  Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,
  Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.
  ESTs of Japanese morning glory
  Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

```

FEATURES

```

source
  Location/Qualifiers
    1..561
      /organism="Ipomoea nil"
      /mol_type="mRNA"
      /cultivar="Tokyo-koeki standard"
      /db_xref="taxon:35883"
      /clone="jmi8h18"
      /tissue type="mixture of flower and flower bud"
      /clone_lib="Ipomoea nil mixture of flower and flower bud"

```

ORIGIN

```

Query Match      83.6%; Score 18.4; DB 4; Length 561;
Best Local Similarity 95.0%; Pred. No. 2.2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY

```

3 CCTTCCTTCATCCTCAGAC 22
  |||||
Db  498 CCTTCCTTCATCCTCAGAC 479

```

RESULT 33

```

BJ554774/c
LOCUS
DEFINITION
  BJ554774
  CDNA clone jmi5d20 5', mRNA sequence.
ACCESSION
  BJ554774
VERSION
  BJ554774.1 GI:27236594
KEYWORDS
  EST.
SOURCE
  Ipomoea nil (Japanese morning glory)
ORGANISM
  Ipomoea nil
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  asterids; lamids; Solanales; Convolvulaceae; Ipomoeae; Ipomoea.
  1 (bases 1 to 566)
REFERENCE
  Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,
  Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.
  ESTs of Japanese morning glory
  Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856

```

```

Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
source
1. .566
/organism="Ipomoea nil"
/mol type="rRNA"
/cultivar="Tokyo-kokei standard"
/db_xref="taxon:35883"
/clone="jml5d20"
/tissue type="mixture of flower and flower bud"
/clone_lib=Ipomoea nil mixture of flower and flower bud"

ORIGIN
Query Match 83.6%; Score 18.4; DB 4; Length 566;
Best Local Similarity 95.0%; Pred. No. 2.2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 CCTTCCTTCATCCTCAGAC 22
|||||
Db 499 CCTTCCTTCATCCTCAGAC 480

RESULT 34
BJ564239/c
LOCUS
DEFINITION BJ564239 Ipomoea nil mixture of flower and flower bud Ipomoea nil
cDNA clone jmk412 5', mRNA sequence.
ACCESSION BJ564239.1 GI:27246059
VERSION
KEYWORDS
SOURCE
ORGANISM Ipomoea nil (Japanese morning glory)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Convolvulaceae; Ipomoeae; Ipomoea.
REFERENCE
Hoshino,A., Seki,M., Shin-i,T., Carninci,P., Kamiya,A., Shiraki,T.,
Nitasaka,E., Shinozaki,K., Hayashizaki,Y., Kohara,Y. and Iida,S.
ESTs of Japanese morning glory
JOURNAL
Unpublished (2002)
COMMENT
Contact: Tadasu Shin-i
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
source
1. .569
/organism="Ipomoea nil"
/mol type="rRNA"
/cultivar="Tokyo-kokei standard"
/db_xref="taxon:35883"
/clone="jmk412"
/tissue type="mixture of flower and flower bud"
/clone_lib=Ipomoea nil mixture of flower and flower bud"

ORIGIN
Query Match 83.6%; Score 18.4; DB 4; Length 569;
Best Local Similarity 95.0%; Pred. No. 2.2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 CCTTCCTTCATCCTCAGAC 22
|||||
Db 499 CCTTCCTTCATCCTCAGAC 480

RESULT 35
CL160038/c
LOCUS
DEFINITION CL160038 586 bp DNA linear GSS 06-JAN-2004
104_350_10804692_116_31823_132 Sorghum methylation-filtered library
(LiBD: 104) Sorghum bicolor genomic clone 10804692, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM Sorghum bicolor (sorghum)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE
1 (bases 1 to 586)
Budiman,M.A., Flick,E., Jones,J., Nunberg,A., Citek,R.W.,
Robbins,D., Rohlfing,T., Bradford,K., Fries,J., McManamy,J.,
Trani,L., Iak,A., Zimmerman,C., Lakey,N. and Bedell,J.A.
GeneThresher methylation filtered genomic sequences from Sorghum
bicolor
Unpublished (2004)
JOURNAL
Contact: Bedell JA
COMMENT
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 5975
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: 350 row: f column: 12
Seg primer: T3 Reverse
Class: shotgun
High quality sequence stop: 586.
Location/Qualifiers
source
1. .586
/organism="Sorghum bicolor"
/mol type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="10804692"
/clone_lib="Sorghum methylation-filtered library (LiBD:
104)"
/notes="Organ: leaf; Vector: pBSK(-); Site 1: HincII; Site 2:
HincII; prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to
1.0 kb fraction, ligated into HincII-digested pBSK(-) vector
and electroporated into E. coli cells. This is a
methylation-filtered library."

ORIGIN
Query Match 83.6%; Score 18.4; DB 9; Length 586;
Best Local Similarity 95.0%; Pred. No. 2.2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 CCTTCCTTCATCCTCAGAC 22
|||||
Db 573 CCTTCCTTCATCCTCAGAC 554

RESULT 36
CO099643/c
LOCUS
DEFINITION CO099643 GR_Ea24K10.r GR_Ea Gossypium raimondii cDNA clone GR_Ea24K10.3
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM Gossypium raimondii
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE
1 (bases 1 to 667)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
JOURNAL
Unpublished (2004)
COMMENT
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona

```

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 24 row: K column: 10.

FEATURES

source
Location/Qualifiers
1. .667
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea24K10"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="Vector: pCMV, SPORT-6.1; Site 1: NotI; Site 2: EcoRV; Library made by Invitrogen with RNA supplied by Wendle lab. Directional cloned into NotI-EV. Colonies plated/picked by Agri. More glycerol clones held in -80."

ORIGIN

Query Match 83.6%; Score 18.4; DB 7; Length 667;
Best Local Similarity 95.0%; Pred. No. 2.2e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 ACCTTCCTTCATCCTCAGA 21
|||
Db 177 ACTTCCTTCATCCTCAGA 158
|||

RESULT 37

CL180445/c
LOCUS
DEFINITION
104,390,10896020,116,31930,164 Sorghum methylation-filtered library (LibID: 104) Sorghum bicolor genomic clone 10896020, genomic survey sequence.

ACCESSION

CL180445
CL180445.1 GI:40692968

KEYWORDS

Sorghum bicolor (sorghum)

SOURCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE

AUTHORS
Budinman, M.A., Flick, E., Jones, J., Nunberg, A., Citek, R.W., Robbins, D., Rohlfing, T., Bradford, K., Fries, J., McMenamy, J., Trani, L., Isak, A., Zimmerman, C., Lakey, N. and Bedell, J.A.
GeneThresher methylation filtered genomic sequences from Sorghum bicolor

JOURNAL

COMMENT
Unpublished (2004)

Contact: Bedell JA

Orion Genomics, LLC

4041 Forest Park Ave, St. Louis, MO 63108, USA

Tel: 314 615 6979

Fax: 314 615 5975

Email: jbedell@oriongenomics.com

Plate: 390 row: 9 column: 20

Seq primer: T3 Reverse

Class: shotgun

High quality sequence stop: 713.

FEATURES

source
Location/Qualifiers
1. .713
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="Atx623"
/db_xref="taxon:4558"
/clone="10896020"
/clone_lib="Sorghum methylation-filtered library (LibID: 104)"
/note="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA prepared from purified nuclei was randomly sheared, end-repaired, size fractionated to enrich for the 0.5 to 5

ORIGIN

Query Match 83.6%; Score 18.4; DB 9; Length 713;
Best Local Similarity 95.0%; Pred. No. 2.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTTCCTTCATCCTCAGAC 22
|||||
Db 571 CCTTCCTTCCTCCTCAGAC 552
|||||

RESULT 38

CO146781
LOCUS
DEFINITION
EST21834 Aspergillus flavus Normalized cDNA Expression Library Aspergillus flavus cDNA clone NAGC050 5' end, mRNA sequence.

ACCESSION

CO146781.1 GI:48900782

VERSION

EST

KEYWORDS

Aspergillus flavus

SOURCE

ORGANISM

Aspergillus flavus

Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;

Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

1 (bases 1 to 733)

AUTHORS

Yu, J., Whitelaw, C.A., Nierman, W.C., Bhatnagar, D. and Cleveland, T.E.

Aspergillus flavus expressed sequence tags for identification of

genes with putative roles in aflatoxin contamination of crops

FEMS Microbiol. Lett. (2004) In press

CONTACT: Yu J

Food and Feed Safety Research Unit

USDA/ARS, Southern Regional Research Center

1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA

Tel: 504 286 4405

Fax: 504 286 4419

Email: jiyu@rrrc.ars.usda.gov

Contact Dr. Yu at USDA/ARS SRRRC (jiyu@rrrc.ars.usda.gov) for clone

information

PCR Primers

FORWARD: M13F

BACKWARD: M13R

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

1. .733

/organism="Aspergillus flavus"

/mol_type="mRNA"

/strain="NRRL 3357"

/db_xref="taxon:5059"

/clone="NAGC050"

/sex="asexual mycelia"

/cell_type="mycelia"

/dev_stage="developmental stages from 18 to 96 hours"

/lab_host="E. coli DH10B T1 resistant cells"

/clone_lib="Aspergillus flavus Normalized cDNA Expression Library"

/note="Vector: pBluescript (SK+) (Stratagene), antibiotic selection marker: Carbenicillin; Site 1: NotI, at the 5 prime end; Site 2: EcoRI, at the 3 prime end; This

normalized cDNA expression library was constructed using a

mixture of mycelial cells grown under eight different

medium conditions and harvested at 5 time points (18, 24,

48, 72, 96 hours). The poly-A sequence was trimmed off

before ligating to vector."

ORIGIN

Query Match 83.6%; Score 18.4; DB 7; Length 733;
Best Local Similarity 95.0%; Pred. No. 2.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CACTTCCTTCCTCCTCAG 20

```

Db          404 CACCTTCCTCCACCTCAG 423
|||||
RESULT 39
BZ936415
LOCUS
DEFINITION
736 bp DNA linear GSS 13-JUN-2003
CH240_87E17.TV CHORI-240 Bos taurus genomic clone CH240_87E17,
genomic survey sequence.
ACCESSION
BZ936415
VERSION
BZ936415.1 GI:31721798
KEYWORDS
GSS.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
REFERENCE
1 (bases 1 to 736)
Larkin,D.M., Everts-van der Wind,A., Rebeiz,M., Schweitzer,P.,
Bachman,S., Green,S., Campos,E.J., Benson,L.D., Edwards,J., Liu,L.,
Womack,J.E., de Jong,P.J. and Lewin,H.A.
A Cattle-Human Comparative Map Built with Cattle BAC-ends and Human
Genome Sequence
JOURNAL
Unpublished (2003)
COMMENT
Other_GSSs: CH240_87E17.TJ
Contact: Harris Lewin
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu
Clones are derived from the bovine BAC library CHORI-240
(http://www.chori.org/bacpac/bovine240.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/ordering/information.htm). This work
was undertaken as part of the International Bovine BAC Mapping
Consortium (IBBMC) by by University of Illinois at Urbana
Champaign, USA with funds provided by grant No. AG202-34480-11828
from USDA-CSREES and AG99-35205-8534 from USDA/NRI (Livestock
Genome Sequencing Initiative)
Plate: 87 row: E column: 17
Seq primer: T7
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..736
/organism="Bos taurus"
/mol_type="genomic DNA"
/strain="breed: Hereford"
/db_xref="taxon:9913"
/clone="CH240_87E17"
/sex="Male"
/cell_type="Blood"
/clone_lib="CHORI-240"
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
library (Male) produced by Pieter de Jong"
ORIGIN
Query Match 83.6%; Score 18.4; DB 8; Length 736;
Best Local Similarity 95.0%; Pred. No. 2.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTTCCTTCCTCCCTCAG 22
|||||
DB 451 CCTTCCTTCCTCCCTCAG 470
|||||

RESULT 40
CL180446
LOCUS
DEFINITION
755 bp DNA linear GSS 06-JAN-2004
CL180446
LOCUS
DEFINITION
104 390 10896020 148 31929 164 Sorghum methylation-filtered library
(LibID: 104) Sorghum bicolor genomic clone 10896020, genomic survey
sequence.
ACCESSION
CL180446
VERSION
CL180446.1 GI:40692969
KEYWORDS
GSS.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
REFERENCE
1 (bases 1 to 755)
Rudiman,M.A., Flick,E., Jones,J., Nunberg,A., Citek,R.W.,
Robbins,D., Rohlfing,T., Bradford,K., Pries,J., McMenamy,J.,
Trani,L., Isak,A., Zimmerman,C., Lakey,N. and Bedell,J.A.
GeneThresher methylation filtered genomic sequences from Sorghum
bicolor
JOURNAL
Unpublished (2004)
COMMENT
Contact: Bedell JA
Orton Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@ortongenomics.com
Plate: 390 row: G column: 20
Seq primer: SWfor Forward
Class: shotgun
High quality sequence stop: 755.
FEATURES
source
Location/Qualifiers
1..755
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/cultivar="ATx623"
/db_xref="taxon:4558"
/clone="10896020"
/clone_lib="Sorghum methylation-filtered library (LibID:
104)"
/note="Organ: leaf; Vector: pBCSK(-); Site_1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to 5
kb fraction, ligated into HincII-digested pBCSK(-) vector
and electroporated into E. coli cells. This is a
methylation-filtered library."
ORIGIN
Query Match 83.6%; Score 18.4; DB 9; Length 755;
Best Local Similarity 95.0%; Pred. No. 2.3e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCTTCCTTCCTCCCTCAG 22
|||||
DB 562 CCTTCCTTCCTCCCTCAG 581
|||||

Search completed: June 4, 2005, 11:46:38
Job time : 1346.17 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 04:22:34 ; Search time 151.598 Seconds
(without alignments)
780.981 Million cell updates/sec

Title: US-09-674-277-23

Perfect score: 20

Sequence: 1 atccccagcgcgtccagctg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_16Dec04:*
2: Geneseqn19808:*
3: Geneseqn19908:*
4: Geneseqn20008:*
5: Geneseqn2001as:*
6: Geneseqn2001bs:*
7: Geneseqn2002as:*
8: Geneseqn2002bs:*
9: Geneseqn2003as:*
10: Geneseqn2003bs:*
11: Geneseqn2003cs:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	3	Aaz36123 Primer de
2	20	100.0	1181	3	Aaz36102 Nucleic a
3	17.4	87.0	3555	13	Adr08248 Full leng
4	16.8	84.0	322	4	Aas40833 cDNA enco
5	16.8	84.0	322	4	Aba06647 Human cDN
6	16.8	84.0	322	6	Abv83984 Human pol
7	16.8	84.0	1545	6	Abt05591 Mycobacte
8	16.8	84.0	2259	5	Aas85920 DNA encod
9	16.8	84.0	3957	4	Aba06411 Human cDN
10	16.8	84.0	3957	4	Abv83748 Human pol
11	16.8	84.0	4135	12	Adg63028 Novel hum
12	16.8	84.0	4309	12	Adg22366 Human sof
13	16.8	84.0	4716	13	Acn38575 Tumour-as
14	16.8	84.0	4816	10	Adf74206 Human nov
15	16.8	84.0	4925	13	Adso9855 Human the
16	16.8	84.0	10437	4	Aas41709 Genomid s
17	16.8	84.0	40304	9	Ada03014 Human NCF
18	16.8	84.0	40304	10	Adb72752 Human NCF
19	16.8	84.0	40304	10	Adc85494 Human NCF
20	16.8	84.0	40304	12	Adm74609 Human car

C 21	16.8	84.0	86574	6	ABK83560	Abk83560 Human cDN
C 22	16.8	84.0	86574	13	ADR52822	Adr52822 Drug ther
C 23	16.8	84.0	110000	4	AAI99682_01	Continuation (2 of
C 24	16.8	84.0	110000	4	AAI99682_02	Continuation (3 of
C 25	16.8	84.0	110000	4	AAI99683_01	Continuation (2 of
C 26	16.8	84.0	110000	4	AAI99683_02	Continuation (3 of
C 27	16.4	82.0	18	10	ADE13446	Adel3446 HLA class
C 28	16.4	82.0	18	10	ADE77633	Adel77633 Human pro
C 29	16.4	82.0	18	12	ADL09296	Adl09296 HLA locus
C 30	16.4	82.0	1155	8	ABX63562	Abx63562 Human cDN
C 31	16.4	82.0	3357	8	AAI63981	AAI63981 Human pol
C 32	16.4	82.0	3357	8	ABZ74567	Abz74567 Secreted
C 33	16.4	82.0	3357	8	ADA98951	Ada98951 Human sec
C 34	16.4	82.0	3357	10	ABZ68099	Abz68099 Human sec
C 35	16.4	82.0	3357	12	ADM24696	Adm24696 Human PRO
C 36	16.4	82.0	4014	8	ACC79241	Acc79241 Human AVA
C 37	16.4	82.0	42373	10	ADC87696	Adc87696 Human man
C 38	16.4	82.0	110000	11	ACN44150_3	Continuation (4 of
C 39	16.4	82.0	135638	10	ABX34289	Abx34289 S. atrool
C 40	15.8	79.0	412	4	AAI92608	AAI92608 Human pol
C 41	15.8	79.0	435	4	AAI90099	AAI90099 Human pol
C 42	15.8	79.0	523	12	ADL35809	Adl35809 Human SER
C 43	15.8	79.0	707	4	AAH06541	Aah06541 Human cDN
C 44	15.8	79.0	848	8	ACA49275	Aca49275 Prokaryot
C 45	15.8	79.0	987	8	ACA37997	Aca37997 Prokaryot

ALIGNMENTS

RESULT 1

AAZ36123

ID AAZ36123 standard; DNA; 20 BP.

XX

AC AAZ36123;

DT 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

DE Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

KW IS91, virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

XX WO9955908-A2.

PN 04-NOV-1999.

PD 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98PR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PI Frechon DTM, Laure FC, Thierry D;

DR WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

PT Escherichia coli, particularly serotype O157:H7, used for detecting these

PT bacteria in food.

PS Claim 5; Page 27; 48pp; French.

CC AAZ36103-27 represent fragments derived from nucleic acid sequences

CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are

CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to

CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and

CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).

CC The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AA236102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies

XX SQ Sequence 20 BP; 3 A; 9 C; 5 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
|||||
Db 1 ATCCAGCGCGCTCCAGCTG 20

RESULT 2

AA236102

ID AA236102 standard; DNA; 1181 BP.

XX AC AA236102;

XX DT 11-FEB-2000 (first entry)

XX Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX DE Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

XX KW enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.

XX OS Escherichia coli.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX DR WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.

XX Claim 1; Fig 2; 48pp; French.

XX The present sequence is specific to enterohemorrhagic Escherichia coli
CC (EHEC). The sequence associated with the presence of virulence factors
CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
CC 68% homology with the virK gene which codes for virulence proteins of
CC Shigella flexneri. The present sequence is of plasmid origin. Fragments
CC of the present sequence are used, as probes and primers, for detection of
CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
CC animal samples, foods or the environment. The fragments are also useful
CC for epidemiological studies

XX SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 1181;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
|||||
Db 41 ATCCAGCGCGCTCCAGCTG 60

RESULT 3

ADR08248/c

ID ADR08248 standard; cDNA; 3555 BP.

XX AC ADR08248;

XX DT 04-NOV-2004 (first entry)

XX Full length human cDNA useful for treating neurological disease Seq 1754.

XX Gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;

XX KW osteoporosis; neurological disease; Alzheimer's disease;

XX KW Parkinson's disease; dementia; short memory; cancer;

XX KW sense of motor function; emotional reaction; fear response; panic;

XX KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;

XX KW tranquiliser.

XX OS Homo sapiens.

XX PN EP1447413-A2.

XX PD 18-AUG-2004.

XX PF 12-FEB-2004; 2004EP-00003145.

XX PR 14-FEB-2003; 2003JP-00102207.

XX PR 09-MAY-2003; 2003JP-00131452.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;

XX PI Wakamatsu A, Ishii S, Nagai K, Irie R;

XX DR WPI; 2004-583265/57.

XX DR P-PSDB; ADR10204.

XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 1754; 2686pp; English.

XX This invention relates to novel, isolated full length human cDNA
CC molecules and the encoded proteins thereof. Specifically, it refers to
CC cDNA clones obtained by an oligo-capping method, where none of these
CC clones are identical to any known human mRNAs. The present invention
CC describes an immunoassay to identify agonists and antagonists, as well as
CC antibodies, antisense molecules and siRNAs that can all be used to bind
CC to and modulate expression of the cDNA molecules. As such, these
CC molecules are useful for diagnostic markers or therapeutic targets for
CC the various diseases or morbid states. In particular, they are useful in
CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
CC disease, Parkinson's disease, dementia, short memory and various cancers,
CC as well as for maintaining equilibrium of sense or motor function, and
CC for treating emotional reaction, fear response and panic. Accordingly,
CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
CC cytostatic and tranquiliser activities. This polynucleotide is a full
CC length human cDNA sequence of the invention. NOTE: This sequence is not
CC given in the sequence listing of the specification but can be obtained on
CC CD-ROM from the European Patent Office, Vienna Sub-office.

XX SQ Sequence 3555 BP; 869 A; 979 C; 948 G; 759 T; 0 U; 0 Other;

Query Match 87.0%; Score 17.4; DB 13; Length 3555;
Best Local Similarity 94.7%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCT 19
|||||
Db 689 ATCCAGCTCGCTCCAGCT 671

RESULT 4
AAS40833
ID AAS40833 standard; cDNA; 322 BP.
XX
AC AAS40833;
XX
DT 17-DEC-2001 (first entry)
XX
DE cDNA encoding novel human enzyme polypeptide #49.
XX
KW Human; oxidoreductase enzymes; transferase; hydrolase; lyase; isomerase;
KW ligase; hyperproliferative disorder; immunodeficiency disorder;
KW autoimmune disorder; neurological disorder; metabolic disorder;
KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW anti arthritic; nephrotropic; anticoagulant; ss.
XX
OS Homo sapiens.
XX
PN WO20015301-A2.
XX
PD 02-AUG-2001.
XX
PP 17-JAN-2001; 2001WO-US001239.
XX
31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.

```
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465566/50.
XX P-PSDB; AAU22963.
XX
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
XX treating neural, immune system, muscular, reproductive, pulmonary,
XX cardiovascular, renal, proliferative disorders and cancerous diseases.
XX
XX Claim 4; SEQ ID NO 59; 1180pp; English.
XX
XX The present invention relates to the isolation of novel human enzyme
XX polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
XX encoding them. The enzyme polypeptides of the invention may comprise the
XX functional classes of oxidoreductases, transferases, hydrolases, lyases,
XX isomerases or ligases. The sequences of the invention are useful in the
XX diagnosis, treatment, prevention and/or prognosis of a wide range of
XX disorders including hyperproliferative disorders (e.g. cancer),
XX immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
XX arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
XX disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
XX cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
XX (e.g. haemophilia), reproductive disorders (e.g. infertility) and
XX infectious disorders (e.g. Influenza). The polynucleotides of the
XX invention can also be used in gene therapy. AAS40785-AAS41684 represent
XX cDNA sequences encoding for the novel human enzyme polypeptides of the
XX invention. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 322 BP; 61 A; 107 C; 100 G; 51 T; 0 U; 3 Other;
XX
XX Query Match 84.0%; Score 16.8; DB 4; Length 322;
XX Best Local Similarity 90.0%; Pred. No. 5e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 ATCCGAGCGCGCTCCAGCTG 20
XX
XX Db 278 AGCCGAGCGCGCGCTCCAGCTG 297
XX
XX RESULT 5
XX ABA06647
XX ID ABA06647 standard; cDNA; 322 BP.
XX
XX AC ABA06647;
XX
XX DT 10-JAN-2002 (first entry)
XX
XX DE Human cDNA SEQ ID NO: 313.
XX
XX Human; gene therapy; neural disorder; immune system disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX proliferative disorder; inflammation; 88.
```

```
XX OS Homo sapiens.
XX XX WO200154474-A2.
XX PN
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001349.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 21-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
```


PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 08-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
DR WPI: 2002-581727/73.
DR P-PSDB; ABP67012.
XX
XX Novel polypeptide useful for diagnosis, prognosis, prevention, and
PT treatment of immune, hyperproliferative, renal, respiratory, and
PT cardiovascular, reproductive, endocrine, gastrointestinal and
PT neurological disorders.
XX
XX Claim 1; SEQ ID NO 313; 369pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABV83682-ABV84101) and proteins
CC (ABP6710-ABP6712) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 322 BP; 61 A; 107 C; 100 G; 51 T; 0 U; 3 Other;

Query Match 84.0%; Score 16.8; DB 6; Length 322;
Best Local Similarity 90.0%; Pred. No. 5e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCGAGCGCGTCCAGCTG 20
DB 278 AGCCGAGCGCGTCCAGCTG 297
RESULT 7
ABT05591/c
ID ABT05591 standard; DNA; 1545 BP.
XX
AC ABT05591;
XX
DT 24-OCT-2002 (first entry)
XX
DE Mycobacterium tuberculosis gene sequence 8.
XX
KW Gene; ds; vaccine; mycobacterial promoter identification;
KW mycobacterial virulence; mycobacterial infection.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200244406-A2.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-GB005250.
XX
PR 28-NOV-2000; 2000GB-00028966.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Vipond R, Shuttleworth H, Ambrose E, Minton NP;
XX
DR WPI: 2002-566590/60.
DR P-PSDB; ABJ04673.
XX
PT Identifying mycobacterial gene induced/up-regulated during Mycobacterium
PT tuberculosis virulence, by aligning promoter sequence linked to target
PT sequence with sequence of same mycobacterium and identifying target.
XX
XX Claim 8; Page 85-86; 205pp; English.
XX
XX The invention comprises a method of identifying a mycobacterial promoter
CC sequence which is induced or up-regulated during mycobacterial virulence.
CC The method involves identifying a mycobacterial promoter sequence which
CC is up-regulated during infection of a macrophage by a Mycobacterium
CC tuberculosis host cell that contains a promoter operably linked to a
CC coding sequence of a reporter gene. The method of the invention is useful
CC in the manufacture of a medicament for treating/preventing a
CC mycobacterial infection. The method of the invention is also useful in
CC the manufacture of a diagnostic reagent for identifying a mycobacterial
CC infection. The present DNA sequence represents a Mycobacterium
CC tuberculosis gene sequence that was used in the invention
XX
SQ Sequence 1545 BP; 297 A; 484 C; 476 G; 288 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 6; Length 1545;
Best Local Similarity 90.0%; Pred. No. 5.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCGAGCGCGTCCAGCTG 20
DB 1236 ATCCGAGCGCGTCCAGCTG 1217
RESULT 8
AAS85920/c
ID AAS85920 standard; cDNA; 2259 BP.
XX
AC AAS85920;
XX
DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #21724.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
KW Homo sapiens.
XX
XX WO200175067-A2.
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI P-PSDB; ABG21733.
XX
XX WPI; 2001-639362/73.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 21724; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensic, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2259 BP; 513 A; 659 C; 620 G; 467 T; 0 U; 0 Other;

Query Match 84.0%; Score 16.8; DB 5; Length 2259;
Best Local Similarity 90.0%; Pred. No. 5.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1605 ATCCGAGCGGCTAGCTG 1586

RESULT 9
ABR06411/c
ID ABR06411 standard; cDNA; 3957 BP.
XX
XX ABR06411;
XX
XX 10-JAN-2002 (first entry)
DT
XX Human cDNA SEQ ID NO: 77.
DE

XX Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200154474-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001349.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226581P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.

```
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234988P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246179P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-476161/51.
XX P-PSDB; ABB10189.
XX
XX Isolated nucleic acid molecule encoding an inflammation-associated
XX polypeptide is used in preventing, treating or ameliorating a medical
XX condition.
XX
XX Claim 1; SEQ ID NO 77; 859pp + Sequence Listing; English.
XX
XX The present invention provides human cDNAs, proteins and related genomic
XX DNAs. These can be used in the treatment of neural, immune system,
XX muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
XX renal and proliferative disorders and inflammation. The present sequence
XX is a cDNA of the invention
XX
XX Sequence 3957 BP; 755 A; 1345 C; 1096 G; 761 T; 0 U; 0 Other;
XX
XX Query Match 84.0%; Score 16.8; DB 4; Length 3957;
XX Best Local Similarity 90.0%; Pred. No. 5.4e+02;
XX Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCGAGCGCGTCCAGCTG 20
Db 36 AGCCGAGCGCGCCCCAGCTG 17

RESULT 10
ABV83748/c
ID ABV83748 standard; cDNA; 3957 BP.
XX
XX AC ABV83748;
XX
XX DT 09-DEC-2002 (first entry)
XX
XX Human polynucleotide SEQ ID NO 77.
XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;
XX antiparkinsonian; antiscickling; antianaemic; antiarthritic; cancer;
XX antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine;
XX gene; ss.
XX
XX Homo sapiens.
XX
XX US2002090672-A1.
XX
XX 11-JUL-2002.
XX
XX PF 17-JAN-2001; 2001US-00764853.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 28-JUN-2000; 2000US-0214886P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
```

PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-681727/73.
DR P-PSDB; ABP66776.
XX
PT Novel polypeptide useful for diagnosis, prognosis, prevention, and
PT treatment of immune, hyperproliferative, renal, respiratory,
PT cardiovascular, reproductive, endocrine, gastrointestinal and
PT neurological disorders.
XX
XX Claim 1; SEQ ID NO 77; 369pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABV83682-ABV84101) and proteins
CC (ABP66710-ABP67129) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 3957 BP; 755 A; 1345 C; 1096 G; 761 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 6; Length 3957;
Best Local Similarity 90.0%; Pred. No. 5.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCAGCGCGCTCCAGCTG 20
Db 36 AGCCAGCGCGCGCCAGCTG 17
RESULT 11
ADQ63028/c
ID ADQ63028 standard; cDNA; 4135 BP.
XX AC ADQ63028;
XX DT 07-OCT-2004 (first entry)
XX DE Novel human cDNA sequence #189.
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW cancer.
XX OS Homo sapiens.
XX PN EP1440981-A2.
XX PD 28-JUL-2004.
XX PF 21-JAN-2004; 2004EP-00001196.
XX PR 21-JAN-2003; 2003JP-00102206.
XX PR 09-MAY-2003; 2003JP-00131392.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Nagai K, Irie R;
XX WPI; 2004-535376/52.
XX P-PSDB; ADQ65216.
XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX Claim 1; SEQ ID NO 189; 2449pp; English.
XX The invention relates to 2495 novel polynucleotides (I) and their encoded
CC polypeptides, sequences hybridizing to these nucleotides, sequences
CC encoding partial polypeptides and sequences having 70% or 90% identity to
CC the nucleotide and protein sequences. The nucleotides and polypeptides
CC are useful as diagnostic markers or therapeutic target for the diseases
CC or morbid states. They are also useful for treating osteoporosis,
CC neurological diseases, Alzheimer's diseases, Parkinson's disease,
CC dementia and various cancers. This sequence corresponds to a nucleotide
CC sequence of the invention.
XX
SQ Sequence 4135 BP; 674 A; 1448 C; 1206 G; 807 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 12; Length 4135;
Best Local Similarity 90.0%; Pred. No. 5.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCAGCGCGCTCCAGCTG 20
Db 405 AGCCAGCGCGCGCCAGCTG 386
RESULT 12

```
ADQ22366/c
ID ADQ22366 standard; DNA; 4309 BP.
AC ADQ22366;
XX
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5186.
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
XX Homo sapiens.
XX
XX WO2004048938-A2.
XX
XX 10-JUN-2004.
XX
XX 26-NOV-2003; 2003WO-US038193.
XX
XX 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
XX
XX WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 5186; 210pp; English.
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
XX Sequence 4309 BP; 750 A; 1472 C; 1258 G; 829 T; 0 U; 0 Other;
SQ
Query Match 84.0%; Score 16.8; DB 12; Length 4309;
Best Local Similarity 90.0%; Pred. No. 5.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCCAGCGCGCTCCAGCTG 20
| | | | | | | | | | | | | | | |
Db 270 AGCCACGCGCGCCCCAGCTG 251

RESULT 13
ACN38575/c
ID ACN38575 standard; cDNA; 4716 BP.
XX
XX ACN38575;
AC
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Tumour-associated antigenic target (TAT) cDNA DNA324893, SEQ ID NO:2170.
DE
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;

ADQ22366/c
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
XX Homo sapiens.
OS
XX WO2004030615-A2.
XX
XX 15-APR-2004.
PD
XX 29-SEP-2003; 2003WO-US028547.
XX
XX 02-OCT-2002; 2002US-0414971P.
XX
XX (GETH ) GENENTECH INC.
XX
XX Wu TD, Zhang Z, Zhou Y;
PI
XX WPI; 2004-347921/32.
XX
XX P-PSDB; ABM80839.
DR
XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
XX Claim 1; SEQ ID NO 2170; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
XX Sequence 4716 BP; 772 A; 1652 C; 1415 G; 877 T; 0 U; 0 Other;
SQ
Query Match 84.0%; Score 16.8; DB 13; Length 4716;
Best Local Similarity 90.0%; Pred. No. 5.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCCAGCGCGCTCCAGCTG 20
| | | | | | | | | | | | | | | |
Db 858 AGCCACGCGCGCCCCAGCTG 839

RESULT 14
ADP74206/c
ID ADP74206 standard; cDNA; 4816 BP.
XX
XX ADP74206;
AC
XX
XX 26-FEB-2004 (first entry)
DT
XX
XX Human novel brain/hippocampus cDNA #24.
DE
XX Human; brain; hippocampus; gene therapy; mental illness;
KW proteome analysis; ss; gene.
```


XX OS Homo sapiens.
XX PN JP2003009886-A.
XX PD 14-JAN-2003.
XX PP 21-DEC-2001; 2001JP-00390441.
XX PR 22-DEC-2000; 2000JP-00389742.
XX PR 29-MAR-2001; 2001JP-00095524.
XX PR 25-APR-2001; 2001JP-00127066.
XX XX
XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.
XX PA (PROT-) PROTEIN EXPRESS KK.
XX XX
XX WPI; 2003-735084/70.
XX P-PSDB; ADF74136.
XX
XX Novel isolated polypeptide useful for screening compounds which modulate
XX PT the activity of polypeptide and thus have therapeutic applications.
XX
XX Disclosure; SEQ ID NO 94; 319pp; Japanese.
XX
XX The invention relates to an isolated human polypeptide consisting of a
XX CC sequence of any one of 70 fully defined sequences (S1) appearing as
XX CC ADF74113-ADF74182 being encoded by one of 70 disclosed cDNA sequences
XX CC isolated from brain or hippocampus tissue, or a sequence of (S1) having
XX CC deletion, substitution, or addition and essentially retaining the same
XX CC biological activity of (S1). Also included are DNA that encodes the
XX CC protein, DNA which hybridises with the encoding DNA and codes a
XX CC polypeptide which substantially as the same activity as that of a
XX CC polypeptide above, a gene including the DNA, a recombinant polypeptide
XX CC encoded by the gene, an antibody against the protein, DNA chip which
XX CC comprises the DNA, a polypeptide chip comprising the protein, an antibody
XX CC chip comprising the antibody, a recombinant vector containing the DNA, a
XX CC transformed host cell containing the vector, an antisense polynucleotide
XX CC which has a sequence complementary to the DNA, a kit for screening
XX CC compounds that interact with the protein, test compounds which interact
XX CC with the protein and homologous proteins having 70% sequence identity
XX CC with (S1). The protein is useful for identifying compounds which interact
XX CC with the protein e.g., compounds which inhibit or activate the protein.
XX CC Compounds that inhibit the protein activity are useful as therapeutic
XX CC agents for treating disorders characterised by aberrant protein activity.
XX CC The DNA is useful for producing the protein by recombinant techniques, as
XX CC a probe in diagnostic techniques for detecting abnormalities in the gene,
XX CC and in gene therapy techniques. The DNA is useful for creating animal
XX CC models of disease. The DNA chips are useful for diagnosing any
XX CC abnormalities of the DNA in biological samples obtained from human
XX CC subjects, where the abnormalities result in mental illness. The antibody
XX CC is useful for detecting the protein in biological samples. The protein
XX CC chip is useful for functional analysis of the protein, e.g. expression of
XX CC the protein, interaction of the protein, post-translational modification
XX CC of the protein. The antibody chip is useful for analysing disease states
XX CC associated with polypeptide. The antibody chip is also useful in proteome
XX CC analysis. The present sequence encodes a human brain/hippocampus protein
XX CC of the invention.
XX
XX SQ Sequence 4816 BP; 781 A; 1682 C; 1450 G; 903 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 10; Length 4816;
Best Local Similarity 90.0%; Pred. No. 5.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCGAGCGCGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 957 AGCCGAGCGCGCGCCGAGCTG 938
RESULT 15
AD809855/c
ID ADS09855 standard; DNA; 4925 BP.
XX

AC ADS09855;
XX 16-DEC-2004 (first entry)
XX DT Human therapeutic DNA - SEQ ID 92.
XX DE
XX KW antinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
XX KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX OS Homo sapiens.
XX PN WO2004080148-A2.
XX PD 23-SEP-2004.
XX PF 30-SEP-2003; 2003WO-US030720.
XX PR 02-OCT-2002; 2002US-0416186P.
XX XX (NUVE-) NUVELO INC.
XX PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
XX PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Wang G, Zhou P;
XX DR WPI; 2004-668857/65.
XX PT P-PSDB; ADS10539.
XX PT New polynucleotide, useful in preparing a composition for diagnosing or
XX PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
XX PT aplastic anemia or cancer for promoting wound healing.
XX PS Claim 1; SEQ ID NO 92; 718pp; English.
XX CC The invention relates to a novel isolated polynucleotide and the encoded
XX CC polypeptide. The molecules of the invention demonstrate antinflammatory,
XX CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
XX CC be useful in preparing a composition for diagnosing or treating
XX CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
XX CC disorders, such as aplastic anaemia or cancer, as well as for promoting
XX CC wound healing. The molecules may also be utilised during gene therapy
XX CC procedures. The current sequence is that of a human therapeutic DNA of
XX CC the invention. The current sequence is not shown explicitly within the
XX CC specification but can be accessed from the WIPO web-site.
XX SQ Sequence 4925 BP; 806 A; 1691 C; 1463 G; 965 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 13; Length 4925;
Best Local Similarity 90.0%; Pred. No. 5.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCGAGCGCGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1050 AGCCGAGCGCGCCGAGCTG 1031
RESULT 16
AAS41709
ID AAS41709 standard; DNA; 10437 BP.
XX AC AAS41709;
XX DT 17-DEC-2001 (first entry)
XX DE Genomic sequence #25 encoding novel human enzyme polypeptide.
XX KW Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
XX KW ligase; hyperproliferative disorder; immunodeficiency disorder;
XX KW autoimmune disorder; neurological disorder; metabolic disorder;
XX KW inflammatory disorder; cardiovascular disorder; reproductive disorder;
XX KW blood-related disorder; infectious disorder; gene therapy; cytostatic;
XX KW anti arthritic; nephrotropic; anticoagulant; ds.

OS Homo sapiens.
XX WO200155301-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US001239.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184564P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216547P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259676P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA

XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-465566/50.
XX Novel polypeptides and polynucleotides useful for diagnosing, preventing,
PT treating neural, immune system, muscular, reproductive, pulmonary,
PT cardiovascular, renal, proliferative disorders and cancerous diseases.
XX Disclosure; SEQ ID NO 1835; 1180pp; English.
XX The present invention relates to the isolation of novel human enzyme
CC polypeptides (AAU22915-AAU23814), and the cDNA and genomic sequences
CC encoding them. The enzyme polypeptides of the invention may comprise the
CC functional classes of oxidoreductases, transferases, hydrolases, lyases,
CC isomerases or ligases. The sequences of the invention are useful in the
CC diagnosis, treatment, prevention and/or prognosis of a wide range of
CC disorders including hyperproliferative disorders (e.g. cancer),
CC immunodeficiency disorders (e.g. AIDS) autoimmune disorders (e.g.
CC arthritis), neurological disorders (e.g. Alzheimer's disease), metabolic
CC disorders (e.g. phenylketonuria), inflammatory disorders (e.g. asthma),
CC cardiovascular disorders (e.g. atherosclerosis), blood-related disorders
CC (e.g. haemophilia), reproductive disorders (e.g. infertility) and
CC infectious disorders (e.g. influenza). The polynucleotides of the
CC invention can also be used in gene therapy. AAS41685-AAS42192 represent
CC DNA sequences encoding for the novel human enzyme polypeptides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 10437 BP; 2265 A; 3267 C; 2795 G; 2110 T; 0 U; 0 Other;
SQ

Query Match 84.0%; Score 16.8; DB 4; Length 10437;
Best Local Similarity 90.0%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
DB 8835 AGCCGAGCGGCTCCAGCTG 8854

RESULT 17
ADA03014/c
ID ADA03014 standard; DNA; 40304 BP.
XX
AC ADA03014;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human NCF4 carcinoma associated gene, SEQ ID NO:1532.
XX
KW Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003057146-A2.
XX
PD 17-JUL-2003.
XX
PF 26-DEC-2002; 2002WO-US041414.
XX
PR 26-DEC-2001; 2001US-00035832.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW;
XX
DR WPI; 2003-587068/55.
XX
PT New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.

XX Claim 1; SEQ ID NO 1532; 245pp; English.
XX The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed human CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 40304 BP; 9854 A; 9739 C; 9912 G; 10276 T; 0 U; 523 Other;
SQ

Query Match 84.0%; Score 16.8; DB 9; Length 40304;
Best Local Similarity 90.0%; Pred. No. 5.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
DB 20556 AGCCGAGCGGCTCCAGCTG 20537

RESULT 18
ADB72752/c
ID ADB72752 standard; DNA; 40304 BP.
XX
AC ADB72752;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human NCF4 gene.
XX
KW human; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.
XX
OS Homo sapiens.
XX
PN WO2003008583-A2.
XX
PD 30-JAN-2003.
XX
PF 26-DEC-2001; 2001WO-US051291.
XX
PR 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-239337/23.
XX
PT New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.
XX

```
PS Claim 1; SEQ ID NO 580; 2304pp; English.
XX
CC The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The
CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a human gene of the invention.
XX
SQ Sequence 40304 BP; 9854 A; 9739 C; 9912 G; 10276 T; 0 U; 523 Other;

Query Match      84.0%; Score 16.8; DB 10; Length 40304;
Best Local Similarity 90.0%; Pred. No. 5.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
Db 20556 AGCCAGCGCTCCAGCTG 20537

RESULT 19
ADC85494/c
ID ADC85494 standard; DNA; 40304 BP.
XX
AC ADC85494;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human Ncf4 genomic sequence.
XX
KW Cytostatic; gene therapy; vaccine; cancer; carcinoma-associated gene; CA;
KW secreted; transmembrane; intracellular; ds.
XX
OS Homo sapiens.
XX
PN WO2003045230-A2.
XX
PD 05-JUN-2003.
XX
PF 02-DEC-2002; 2002WO-US038582.
XX
PR 30-NOV-2001; 2001US-00997722.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2003-513603/48.
XX
PT New recombinant nucleic acid comprising a nucleotide sequence of any of
PT the carcinoma-associated (CA) genes, useful for screening for drug
PT candidates for diagnosing or treating carcinomas.
XX
PS Claim 1; SEQ ID NO 280; 983pp; English.
XX
CC The invention relates to a recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the fully defined carcinoma-
CC associated (CA) genes from the 50 tables given in the specification. The
CC CA proteins are secreted, transmembrane or intracellular proteins. The
CC recombinant nucleic acids are useful for screening for drug candidates
CC for diagnosing or treating carcinomas. Sequences given in ADC85215-
CC ADC85514 represent CA genes of the invention.
XX
SQ Sequence 40304 BP; 9854 A; 9739 C; 9912 G; 10276 T; 0 U; 523 Other;

Query Match      84.0%; Score 16.8; DB 10; Length 40304;
Best Local Similarity 90.0%; Pred. No. 5.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
Db 20556 AGCCAGCGCTCCAGCTG 20537

RESULT 20
ADM74609/c
ID ADM74609 standard; DNA; 40304 BP.
XX
AC ADM74609;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human carcinoma associated (CA) nucleic acid #139.
XX
KW Human; carcinoma associated nucleic acid; CA nucleic acid; gene; ds;
KW carcinoma associated protein; CAP; carcinoma; leukaemia; lymphoma;
KW cytostatic.
XX
OS Homo sapiens.
XX
PN US2004072154-A1.
XX
PD 15-APR-2004.
XX
PF 30-NOV-2001; 2001US-00997722.
XX
PR 22-DEC-2000; 2000US-00747377.
PR 02-MAR-2001; 2001US-00798586.
XX
PA (MORR/) MORRIS D W.
PA (ENG/) ENGELHARD E K.
XX
PI Morris DW, Engelhard EK;
XX
DR WPI; 2004-328562/30.
XX
PT New carcinoma associated gene or protein, useful for preparing a
PT composition for diagnosing or treating carcinoma e.g., leukemia or
PT lymphoma.
XX
PS Claim 1; SEQ ID NO 280; 29pp; English.
XX
CC The invention relates to new recombinant nucleic acids. The invention
CC also relates to a host cell comprising a recombinant nucleic acid or
CC expression vector, an expression vector comprising a recombinant nucleic
CC acid, a recombinant protein, a method of screening for drug candidates, a
CC method of screening for a bioactive agent capable of binding to a
CC carcinoma associated protein (CAP) encoded by a nucleotide sequence, a
CC method of screening for a bioactive agent capable of modulating the
CC activity of a CAP, a method of evaluating the effect of a candidate
CC carcinoma drug, a method of diagnosing carcinoma, a method for inhibiting
CC the activity of a CAP, a method of treating carcinomas, a method of
CC neutralising the effect of a CAP and a method of diagnosing carcinoma or
CC propensity to carcinoma. A method of evaluating the effect of a candidate
CC carcinoma drug comprises administering the drug to a patient, removing a
CC cell sample from the patient and determining alterations in the
CC expression or activation of a gene comprising the nucleotide sequence. A
CC method of diagnosing carcinoma comprises determining the expression of
CC one or more genes comprising the nucleic acid sequence in a first tissue
CC type of a first individual and comparing the expression of the gene from
CC a second normal tissue type from the first individual or a second
CC unaffected individual, where a difference in the expression indicates
CC that the first individual has carcinoma. A method of inhibiting the
CC activity of a CAP comprises binding an inhibitor to the CAP. Treating
CC carcinomas comprises administering to a patient an inhibitor of CAP.
CC Neutralising the effect of a CAP comprises contacting an agent specific
CC for the CAP. The polypeptide specifically binds to the protein encoded by
CC the nucleic acid. It comprises an antibody that specifically binds to the
CC protein encoded by the nucleic acid. The nucleic acids are useful for
CC preparing a composition for diagnosing or treating carcinoma e.g.,
CC leukaemia or lymphoma. This sequence represents a human carcinoma
CC associated (CA) nucleic acid of the invention. Note: The sequence data
CC for this patent did not form part of the printed specification but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html.
```

XX SQ Sequence 40304 BP; 9854 A; 9739 C; 9912 G; 10276 T; 0 U; 523 Other;
Query Match 84.0%; Score 16.8; DB 12; Length 40304;
Best Local Similarity 90.0%; Pred. No. 5.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCGAGCGGCTCCAGCTG 20
DB 20556 AGCCGAGCGGCTCCAGCTG 20537
RESULT 21
ABK83560/c
ID ABK83560 standard; cDNA; 86574 BP.
XX AC ABK83560;
DT 14-AUG-2002 (first entry)
XX DE Human cDNA differentially expressed in granulocytic cells #131.
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
KW viral infection; parasitic infection; protozoal infection;
KW fungal infection; sterile inflammatory disease; psoriasis;
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;
KW adult respiratory distress syndrome; inflammatory bowel disease;
KW Crohn's disease; ulcerative colitis; periodontal disease;
KW granulocyte activation; chronic inflammation; allergy.
XX OS Homo sapiens.
XX WO200228999-A2.
XX 11-APR-2002.
XX 03-OCT-2001; 2001WO-US030821.
XX 03-OCT-2000; 2000US-0237189P.
XX (GENE-) GENE LOGIC INC.
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX Detecting granulocyte activation by detecting differential expression of
PT genes associated with granulocyte activation, which serves as diagnostic
PT markers that is useful for monitoring disease states and drug toxicity.
XX Claim 1; SEQ ID NO 131; 114pp; English.
XX The invention relates to detecting (M1) granulocyte (GC) activation
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by
CC DNA chip analysis as given in the specification, and comparing the
CC expression level to an expression level in an unactivated GC, where
CC differential expression of Gs is indicative of GCA. Also included are
CC modulating (M2) GA by contacting GC with an agent that alters the
CC expression of at least one gene in Gs; (2) screening (M3) for an agent
CC capable of modulating GCA or an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease using the gene expression
CC profile; (3) detecting (M4) an inflammation (especially chronic) in a
CC tissue, an allergic response in a subject, exposure of a subject to a
CC pathogen or sterile inflammatory disease, by detecting the level of
CC expression in a sample of the tissue of gene(s) from Gs, where the level
CC of expression of the gene is indicative of inflammation; (4) treating
CC (M5) an inflammation (especially chronic) or in a tissue, an allergic
CC response in a subject, exposure of a subject to a pathogen or sterile
CC inflammatory disease, by contacting a tissue having inflammation with an
CC agent that modulates the expression of gene(s) from Gs in the tissue. M1
CC is useful for detecting GCA; M2 is useful for modulating GA; M3 is useful

CC for screening an agent capable of modulating GCA preferably in an
CC inflammation in a tissue; M4 is useful for detecting an inflammation
CC (especially chronic) in a tissue, an allergic response in a subject,
CC exposure of a subject to a pathogen or sterile inflammatory disease (e.g.
CC psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis,
CC cardiac reperfusion injury, renal reperfusion injury, ARDS, adult
CC respiratory distress syndrome, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, periodontal disease; also bacterial
CC infection, viral infection, parasitic infection, protozoal infection,
CC fungal infection and M5 is useful for treating one of the above
CC conditions. The present sequence represents a gene differentially
CC expressed in granulocytes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 86574 BP; 22071 A; 20398 C; 21552 G; 22553 T; 0 U; 0 Other;
Query Match 84.0%; Score 16.8; DB 6; Length 86574;
Best Local Similarity 90.0%; Pred. No. 5.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCGAGCGGCTCCAGCTG 20
DB 19631 AGCCGAGCGGCTCCAGCTG 19612
RESULT 22
ADR52822/c
ID ADR52822 standard; DNA; 86574 BP.
XX AC ADR52822;
XX 18-NOV-2004 (first entry)
XX Drug therapy altered expressed gene #173.
XX drug activity monitoring; expression profile; gene expression;
KW peripheral blood sample; peripheral blood mononuclear cell; drug therapy;
KW CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;
KW mTOR; da.
XX OS Homo sapiens.
XX WO2004072265-A2.
XX 26-AUG-2004.
XX 11-FEB-2004; 2004WO-US004118.
XX 11-FEB-2003; 2003US-0446133P.
XX 03-APR-2003; 2003US-0459782P.
XX 23-JAN-2004; 2004US-0538246P.
XX (AMHP) WYETH.
XX (BURC/) BURCZYNSKI M.
XX (TWIN/) TWINE N.
XX (DORN/) DORNER A J.
XX (TREP/) TREPICCHIO W L.
XX Burczynski M, Twine N, Dorner AJ, Trepicchio WL;
XX WPI; 2004-642301/62.
XX Monitoring drug activities in vivo comprises comparing an expression
PT profile of a gene in a peripheral blood sample of a patient before and
PT after drug therapy.
XX Disclosure; SEQ ID NO 173; 136pp; English.
XX The invention relates to a method of monitoring drug activities in vivo
CC by comparing an expression profile of at least one gene in a peripheral
CC blood sample of a patient to a reference expression profile of the at

WP	AAI99683_11	1100001	1210000
WP	AAI99683_12	1200001	1310000
WP	AAI99683_13	1300001	1410000
WP	AAI99683_14	1400001	1510000
WP	AAI99683_15	1500001	1610000
WP	AAI99683_16	1600001	1710000
WP	AAI99683_17	1700001	1810000
WP	AAI99683_18	1800001	1910000
WP	AAI99683_19	1900001	2010000
WP	AAI99683_20	2000001	2110000
WP	AAI99683_21	2100001	2210000
WP	AAI99683_22	2200001	2310000
WP	AAI99683_23	2300001	2410000
WP	AAI99683_24	2400001	2510000
WP	AAI99683_25	2500001	2610000
WP	AAI99683_26	2600001	2710000
WP	AAI99683_27	2700001	2810000
WP	AAI99683_28	2800001	2910000
WP	AAI99683_29	2900001	3010000
WP	AAI99683_30	3000001	3110000
WP	AAI99683_31	3100001	3210000
WP	AAI99683_32	3200001	3310000
WP	AAI99683_33	3300001	3410000
WP	AAI99683_34	3400001	3510000
WP	AAI99683_35	3500001	3610000
WP	AAI99683_36	3600001	3710000
WP	AAI99683_37	3700001	3810000
WP	AAI99683_38	3800001	3910000
WP	AAI99683_39	3900001	4010000
WP	AAI99683_40	4000001	4110000
WP	AAI99683_41	4100001	4210000
WP	AAI99683_42	4200001	4310000
WP	AAI99683_43	4300001	4403765

Query Match 84.0%; Score 16.8; DB 4; Length 110000;

Best Local Similarity 90.0%; Pred. No. 5.9e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0;

Qy	1	ATCCGAGCGCTCCAGCTG	20
Db	6631	ATCGACGCGCTCCAGGTG	6612

RESULT 27

ADE13446

ID ADE13446 standard; DNA; 18 BP.

AC ADE13446;

XX

XX

DT 29-JAN-2004 (first entry)

XX

DE HLA class I allele specific primer #62.

XX es; primer; PCR; human; Human Leukocyte Antigen; HLA; genotype

XX Homo sapiens.

OS

XX US2003165884-A1.

PN

XX

XX

PD 04-SEP-2003.

XX

XX 25-APR-2002; 2002US-00133779.

XX

XX 20-DEC-1999; 99US-0172768P.

PR

XX 20-DEC-2000; 2000US-00747391.

XX

PA (STEM-) STEMCYTE INC.

XX

XX Chow R, Tonai R;

PI

XX

XX WPI; 2003-874916/81.

DR

XX

XX

PT Identifying class I or II Human Leukocyte Antigen genotypes us

PT hybridization and amplification assays.
XX Claim 7; SEQ ID NO 64; 66pp; English.
XX The invention relates to a method of identifying a class I or II Human
XX Leukocyte Antigen (HLA) genotype of a subject using hybridisation and
XX amplification assay. The method is used for determining the HLA genotype
XX of a subject. The present sequence represents a HLA class I allele
XX specific primer.
XX
SQ Sequence 18 BP; 3 A; 8 C; 4 G; 3 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 10; Length 18;
Best Local Similarity 94.4%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCT 19
Db 1 TACCAGCGCGCTCCAGCT 18

RESULT 28
ADE77633
ID ADE77633 standard; DNA; 18 BP.
XX AC ADE77633;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human probe SB221 for elongation mediated multiplexed analysis of HLA-B.
XX
KW probe; ss; human; CFTR; human leukocyte antigen; HLA; genetic testing;
KW carrier screening; genotyping; profiling; polymorphic;
KW multiplexed elongation assay; enzymatic recognition;
KW cystic fibrosis conductance transmembrane regulator.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX WO2003034029-A2.
XX
PD 24-APR-2003.
XX
PF 15-OCT-2002; 2002WO-US033012.
XX
PR 15-OCT-2001; 2001US-0329427P.
XX
PR 15-OCT-2001; 2001US-0329428P.
PR 15-OCT-2001; 2001US-0329619P.
PR 15-OCT-2001; 2001US-0329620P.
PR 14-MAR-2002; 2002US-0364416P.
XX
XX (BIOA-) BIOARRAY SOLUTIONS LTD.
XX
XX Li AX, Hashmi G, Seul M;
XX
XX WPI; 2003-393553/37.
XX
XX Concurrent interrogation of a number of polymorphic sites, useful for
XX genetic testing, carrier screening, genetic profiling, and identity
XX testing, comprises conducting a multiplexed elongation assay using
XX probes.
XX
XX Example 9; Page 48; 143pp; English.
XX
XX This invention relates to a novel method for the concurrent interrogation
XX of a number of polymorphic sites in the presence of, and without
XX interference from, non-designated polymorphic sites. Specifically, it
XX comprises conducting a multiplexed elongation assay by applying one or
XX more temperature cycles to achieve linear amplification of the target or
XX a combination of annealing and elongation steps under temperature-
XX controlled conditions. Furthermore, this detection method uses probe
XX extension or elongation and relies on enzymatic recognition, a superior
XX technique that no longer depends on differential hybridisation. The

CC present invention describes probes and methods useful for identifying or
CC detecting polymorphisms at one or more designated sites, such that they
CC can identify mutations within the cystic fibrosis conductance
CC transmembrane regulator (CFTR) or the human leukocyte antigen (HLA)
CC genes. In addition, concurrent interrogation of a multiplicity of
CC polymorphic sites is useful for genetic testing, carrier screening,
CC genotyping or genetic profiling, and identity testing. This
CC oligonucleotide is a human probe used for the elongation mediated
CC multiplexed analysis of HLA-B, in an exemplification of the invention.
XX
SQ Sequence 18 BP; 3 A; 8 C; 4 G; 3 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 10; Length 18;
Best Local Similarity 94.4%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCT 19
Db 1 TACCAGCGCGCTCCAGCT 18

RESULT 29
ADL09296
ID ADL09296 standard; DNA; 18 BP.
XX AC ADL09296;
XX
DT 06-MAY-2004 (first entry)
XX
DE HLA locus-specific capture oligonucleotide #62.
XX
KW ss; primer; human leukocyte antigen; HLA; HLA genotyping; human; PCR.
XX
OS Homo sapiens.
XX
XX US6670124-B1.
XX
XX 30-DEC-2003.
XX
XX 20-DEC-2000; 2000US-00747391.
XX
XX 20-DEC-1999; 99US-0172768P.
XX
XX (STEM-) STEMCYTE INC.
XX
XX Chow R, Tonai R;
XX
XX WPI; 2004-068584/07.
XX
XX Identifying an HLA genotype of a subject by hybridizing the amplification
XX products with an HLA locus-specific capture oligonucleotide and detecting
XX the detectable complexes to identify the HLA genotype of the subject.
XX
XX Example 1; SEQ ID NO 64; 68pp; English.
XX
XX The invention describes a method of identifying a human leukocyte antigen
XX (HLA) genotype of a subject comprising: obtaining a sample comprising a
XX template nucleic acid from the subject; amplifying the template nucleic
XX acid with HLA allele-specific forward primers and HLA allele-specific
XX reverse primers to form amplification products; hybridising the
XX amplification products with an HLA locus-specific capture oligonucleotide
XX ; and detecting the detectable complexes to identify the HLA genotype of
XX the subject. The present sequence represents one of 276 HLA locus-
XX specific capture oligonucleotides of the invention.
XX
SQ Sequence 18 BP; 3 A; 8 C; 4 G; 3 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 12; Length 18;
Best Local Similarity 94.4%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCT 19
Db 1 TACCAGCGCGCTCCAGCT 18

DT 12-MAY-2003 (first entry)
XX Secreted protein gene 353 genomic fragment HTXDC77, SEQ ID NO:1714.
DE
XX Human; secreted protein; cancer; tumour; hyperproliferative disorder;
KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
KW drug screening; chromosome identification; chromosome mapping;
KW cytostatic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
KW antianaemic; vulnery; gene; ds.
XX Homo sapiens.
OS
XX WO200277013-A2.
PN
XX 03-OCT-2002.
PD
XX 26-MAR-2002; 2002WO-US009370.
XX
XX 27-MAR-2001; 2001US-0278650P.
PR
XX 12-SEP-2001; 2001US-00950082.
PR
XX 12-SEP-2001; 2001US-00950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-040578/03.
XX
XX New human secreted proteins and nucleic acids, useful for detecting or
DR treating cancer or other hyperproliferative disorders, autoimmune
XX disorders, inflammatory disorders, HIV disease, hepatitis or anemia.
PT
XX Disclosure; Page 2301-2302; 2474pp; English.
PS
XX ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted
CC protein genes, and ABP00947-ABP01363 represent the proteins they encode.
CC
CC ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The
CC invention also encompasses antibodies specific for the secreted proteins,
CC the use of the secreted proteins in drug screening and recombinant
CC vectors and host cells comprising a nucleic acid of the invention. The
CC secreted proteins are thought to be involved in biological activities
CC associated with cellular signalling, cellular differentiation, cell
CC migration, prohormone activation and neurotransmitter activity. The
CC secreted proteins, nucleic acids encoding them, antibodies or antibody
CC fragments specific for the secreted proteins, and modulators of protein
CC activity are useful for diagnosing or treating cancers or other
CC hyperproliferative disorders. Additionally, the secreted proteins and
CC their nucleic acids may also be used in the treatment of autoimmune
CC disorders, inflammatory disorders, diseases involving angiogenesis, AIDS
CC (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote
CC wound healing. Nucleic acids of the invention may be used for chromosome
CC identification. Chromosome mapping, in gene therapy, for identifying
CC individuals from minute biological samples, as hybridisation probes, and
CC as molecular weight markers. The present sequence represents a human
CC secreted protein genomic fragment referred to in the disclosure of the
XX invention
XX
XX Sequence 3357 BP; 645 A; 947 C; 1035 G; 730 T; 0 U; 0 Other;
Query Match 82.0%; Score 16.4; DB 8; Length 3357;
Best Local Similarity 94.4%; Pred. No. 8.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TCCAGCGCGCTCCAGCT 19
DB 1059 TACCAGCGCGCTCCAGCT 1042
RESULT 33
ADA98951/c
ID ADA98951 standard; DNA; 3357 BP.
XX

AC ADA98951;
XX 20-NOV-2003 (first entry)
XX Human secreted protein-related DNA sequence #544.
DE human; secreted protein; cardiovascular disorder; arrhythmia;
KW atherosclerosis; stroke; endocarditis; congestive heart failure;
KW rheumatic heart disease; cardiomyopathy; hemorrhoids; varicose veins;
KW migraine; thrombosis; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; renal disorder; proliferative disorder; cancer; ds.
XX
XX Homo sapiens.
OS
XX WO2003004623-A2.
PN
XX 16-JAN-2003.
PD
XX 26-MAR-2002; 2002WO-US009922.
XX
XX 27-MAR-2001; 2001US-0278650P.
PR
XX 12-SEP-2001; 2001US-00950082.
PR
XX 12-SEP-2001; 2001US-00950083.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Rosen CA, Ruben SM;
PI
XX WPI; 2003-247946/24.
XX
XX New human secreted polypeptide and nucleic acid molecules, useful for
DR diagnosing, preventing, prognosticating or treating cardiovascular
XX disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
PT thrombosis).
PT
XX Disclosure; SEQ ID NO 1060; 1572pp; English.
PS
XX The invention comprises the amino acid and coding sequence of human
CC secreted proteins. The DNA and protein sequences of the invention are
CC useful in the treatment of cardiovascular disorders, such as: arrhythmia,
CC atherosclerosis, stroke, endocarditis, congestive heart failure,
CC rheumatic heart disease, cardiomyopathy, hemorrhoids, varicose veins,
CC migraine, or thrombosis. The DNA and protein sequences may also be used
CC for treating or preventing: neural disorders, immune system disorders,
CC muscular disorders, reproductive disorders, gastrointestinal disorders,
CC pulmonary disorders, renal disorders, proliferative disorders and/or
CC cancerous diseases. The present DNA sequence is used in the
CC exemplification of the invention. NOTE: The present sequence is shown on
CC the WIPO website.
XX
XX Sequence 3357 BP; 645 A; 947 C; 1035 G; 730 T; 0 U; 0 Other;
Query Match 82.0%; Score 16.4; DB 8; Length 3357;
Best Local Similarity 94.4%; Pred. No. 8.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TCCAGCGCGCTCCAGCT 19
DB 1059 TACCAGCGCGCTCCAGCT 1042
RESULT 34
ABZ68099/c
ID ABZ68099 standard; DNA; 3357 BP.
XX
XX ABZ68099;
AC
XX 26-MAR-2003 (first entry)
DT
XX Human secreted protein encoding genomic DNA SEQ ID NO 1622.
DE
XX Human; secreted protein; neotropic; neuroprotective; cytostatic;
KW

KW viricide; dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW vulnervary; antibacterial; antiparkinsonian; antiscickling; antihaemic;
KW antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective;
KW antiinflammatory; anti allergic; antidiabetic; antitumor; anticonvulsant;
KW antifungal; antiparasitic; cardiac; immune disorder; infection; vaccine;
KW cardiovascular disorder; neurological disease; nephrotropic;
KW gene therapy; gene; ds.

XX Homo sapiens.

XX WO20027186-A2.

XX 03-OCT-2002.

XX 26-MAR-2002; 2002WO-US009188.

XX 27-MAR-2001; 2001US-027850P.

XX 12-SEP-2001; 2001US-00950082.

XX 12-SEP-2001; 2001US-00950083.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-040583/03.

XX New human secreted proteins encoded by genes contained in cDNA clones
PT (e.g. HCCAC19), useful for preventing, treating or diagnosing e.g. AIDS,
PT multiple sclerosis, herpes virus, leukemia, tick-borne encephalitis or
PT West Nile fever.

XX Disclosure; Page 2252-2253; 2423pp; English.

XX The invention relates to novel human genes (ABZ66891-ABZ68209) and the
CC encoded secreted proteins (ABP9470-ABP9872) useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such as
CC myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g.
CC cerebral anoxia and epilepsy; and (f) infectious diseases such as viral,
CC bacterial, fungal and parasitic infections

XX Sequence 3357 BP; 645 A; 947 C; 1035 G; 730 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 10; Length 3357;
Best Local Similarity 94.4%; Pred. No. 8.2e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCT 19

Db 1059 TACCAGCGCGCTCCAGCT 1042

RESULT 35

ID ADM24696/c

XX ADM24696 standard; cDNA; 3357 BP.

XX ADM24696;

XX 20-MAY-2004 (first entry)

XX Human PRO polynucleotide #179.

XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;

KW cardiovascular disorder; renal disorder; proliferative disorder; cancers;
KW systemic lupus erythematosus; rheumatoid arthritis; multiple sclerosis;
KW thyroiditis; anaemia; Grave's disease; diabetes; hepatitis; asthma;
KW allergy; nephritis; Parkinson's disease; Alzheimer's disease;
KW atherosclerosis; myocardial infarction; AIDS; infection.
XX Homo sapiens.
XX US2004014039-A1.
XX 22-JAN-2004.
XX 31-MAY-2002; 2002US-00158057.
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-019123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.

```
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.

PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764890.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX WPI; 2004-108205/11.
XX
XX Secreted and transmembrane PRO polypeptides and nucleic acids for
XX diagnosing, preventing or treating diseases associated with aberrant
XX expression or activity of the polypeptide, e.g. cancer, asthma, AIDS,
XX Parkinson's disease or diabetes.
XX
XX Claim 4; SEQ ID NO 353; 284pp; English.
XX
XX The invention relates to a secreted and transmembrane PRO polypeptides
XX and the polynucleotides encoding them. The polypeptides and
XX polynucleotides are useful in diagnosing, preventing, prognosing or
XX treating diseases or disorders associated with aberrant expression and/or
XX activity of PRO polypeptides, such as neural disorders, immune system
XX disorders, muscular disorders, reproductive disorders, gastrointestinal
XX disorders, pulmonary disorders, cardiovascular disorders, renal
XX disorders, proliferative disorders and/or cancers. In particular, these
XX diseases are systemic lupus erythematosus, rheumatoid arthritis, multiple
XX sclerosis, thyroiditis, anaemia, Grave's disease, diabetes, hepatitis,
XX asthma, allergies, nephritis, Parkinson's disease, Alzheimer's disease,
XX atherosclerosis, myocardial infarction, AIDS, infections, etc. This
XX sequence represents a human PRO polynucleotide of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 3357 BP; 645 A; 947 C; 1035 G; 730 T; 0 U; 0 Other;

Query Match      82.0%; Score 16.4; DB 12; Length 3357;
Best Local Similarity 94.4%; Pred. No. 8.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 TCCAGCGCGCTCCAGCT 19
        | | | | | | | | | | | | | | | | | |
Db      1059 TACCAGCGCGCTCCAGCT 1042

RESULT 36
ACC79241
ID ACC79241 standard; cDNA; 4014 BP.
XX
AC ACC79241;
XX
DT 31-JUL-2003 (first entry)
XX
DE Human AVATAR encoding cDNA SEQ ID NO:12.
XX
KW Human; MSREBP; sterol regulatory element binding protein; SREBP;
KW antiatherosclerotic; antilipaeamic; antidiabetic; cardiant; diabetes;
KW atherosclerosis; hyperlipidaemia; cardiovascular disease; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..4014
XX     /*tag= a
XX     /partial
XX     /product= "AVATAR"
XX     /note= "no stop codon given"
XX
XX WO2003033656-A2.
XX
XX 24-APR-2003.
```


XX DE S. atroolivaceus leinamycin biosynthesis gene cluster.
XX
KW Leinamycin biosynthesis gene cluster; Lmm; open reading frame; ORF;
KW anti-tumour antibiotic; broad spectrum antimicrobial activity;
KW Gram-positive; Gram-negative bacteria; chemical modification; metabolite;
KW apo-carrier protein; holo-carrier protein; tumour; polyketide;
KW hybrid polypeptide/polyketide metabolite; Lmm production; cytostatic;
KW gene; ds.
XX
XX Streptomyces atroolivaceus.
XX
XX WO200277179-A2.
XX
XX 03-OCT-2002.
XX
XX 22-MAR-2002; 2002WO-US008937.
XX
XX 26-MAR-2001; 2001US-0278935P.
XX
XX (REGC) UNIV CALIFORNIA.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Shen B, Cheng Y, Tang G;
XX
XX WPI; 2003-018907/01.
XX P-PSDB; ABU11341, ABU11342, ABU11343, ABU11344, ABU11345, ABU11346,
XX ABU11347, ABU11348, ABU11349, ABU11350, ABU11351, ABU11352, ABU11353,
XX ABU11354, ABU11355, ABU11356, ABU11357, ABU11358, ABU11359, ABU11360,
XX ABU11361, ABU11362, ABU11363, ABU11364, ABU11365, ABU11366, ABU11367,
XX ABU11368, ABU11369, ABU11370, ABU11371, ABU11372, ABU11373, ABU11374,
XX ABU11375, ABU11376, ABU11377, ABU11378, ABU11379, ABU11380, ABU11381,
XX ABU11382, ABU11383, ABU11384, ABU11385, ABU11386, ABU11387, ABU11388,
XX ABU11389, ABU11390, ABU11391, ABU11392, ABU11393, ABU11394, ABU11395,
XX ABU11396, ABU11397, ABU11398, ABU11399, ABU11400, ABU11401, ABU11402,
XX ABU11403, ABU11404, ABU11405, ABU11406, ABU11407, ABU11408, ABU11409,
XX ABU11410, ABU11411.
XX
XX Novel gene cluster responsible for synthesis of leinamycin in
XX Streptomyces atroolivaceus useful for making various peptide and/or
XX polyketide, and/or hybrid polypeptide/polyketide metabolites.
XX
XX Claim 6; Page 81-127; 185pp; English.
XX
XX The present invention relates to the isolation of the Streptomyces
XX atroolivaceus leinamycin (Lmm) biosynthesis gene cluster containing 71
XX open reading frames (ORFs) (ORFs -35 through -1, ORFs lmmA through lmmZ,
XX and ORFs +1 through +9). Leinamycin is a novel anti-tumour antibiotic
XX produced by several Streptomyces species. It exhibits broad spectrum
XX antimicrobial activity against Gram-positive and Gram-negative bacteria,
XX but not against fungi. The polypeptides encoded by the lmm biosynthesis
XX gene cluster ORFs are useful for chemically modifying a molecule in a
XX host cell. The host cell is a bacterium or eukaryotic cell, including a
XX mammalian, yeast, plant, fungal, or insect cell. The molecule is an
XX endogenous metabolite produced by the host cell or exogenously supplied
XX metabolite, or an amino acid, and the polypeptide is a peptide synthetase
XX or amino transferase. The polypeptides encoded by the lmm gene cluster
XX are useful for converting an apo-carrier protein to a holo-carrier
XX protein. Lmm shows potent antitumour activity in tumour models in vivo.
XX The Lmm gene cluster modules and/or catalytic domains are useful for
XX making various peptide and/or polyketide, and/or hybrid
XX polypeptide/polyketide metabolites. The proteins encoded by the ORFs are
XX useful alone, or in combination with other active domains to modify
XX various target substrates. The lmm gene cluster is useful to upregulate
XX endogenous Lmm production to permit Lmm production in cells and/or to
XX make various modified Lmm. Lmm, its analogue, or other polyketide,
XX peptide or hybrid polyketide/peptide metabolites are useful as
XX therapeutic agents, to treat a number of disorders, depending upon the
XX type of metabolites. The present sequence represents the S. atroolivaceus
XX leinamycin biosynthesis gene cluster
XX
XX Sequence 135638 BP; 18570 A; 49096 C; 49039 G; 18933 T; 0 U; 0 Other;

Query Match 82.0%; Score 16.4; DB 10; Length 135638;
Best Local Similarity 94.4%; Pred. No. 9e+02; 1; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 CCACGGCGCTCCAGCTG 20
DB 23908 CTCAGCGGCTCCAGCTG 23925
RESULT 40
AAI92608/c
ID AAI92608 standard; cDNA; 412 BP.
XX
XX AAI92608;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 12668.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX WO200164835-A2.
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004927.
XX
XX 28-FEB-2000; 2000US-00515126.
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX P-PSDB; AAO12677.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 1; SEQ ID NO 12668; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activity/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 412 BP; 103 A; 80 C; 79 G; 150 T; 0 U; 0 Other;
XX
Query Match 79.0%; Score 15.8; DB 4; Length 412;
Best Local Similarity 89.5%; Pred. No. 1.4e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCCGCGCTCCAGCT 19
DB 401 ATCCCGCGCTCCAGCT 383

Search completed: June 4, 2005, 07:28:56

Job time : 154.598 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 07:08:50 ; Search time 190.651 Seconds
(without alignments)
644.888 Million cell updates/sec

Title: US-09-674-277-23

Perfect score: 20
Sequence: 1 atccagcgctccagctg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	322	9 US-09-764-853-313	Sequence 313, App
2	16.8	84.0	1545	13 US-10-432-934-17	Sequence 17, Appl
3	16.8	84.0	1782	13 US-10-027-633-97445	Sequence 97445, A
4	16.8	84.0	1782	13 US-10-027-633-97446	Sequence 97446, A
5	16.8	84.0	1782	17 US-10-027-633-97445	Sequence 97445, A
6	16.8	84.0	1782	17 US-10-027-633-97446	Sequence 97446, A
7	16.8	84.0	3957	9 US-09-764-853-77	Sequence 77, Appl
8	16.8	84.0	4309	18 US-10-723-860-5186	Sequence 5186, Ap
9	16.8	84.0	40304	11 US-09-997-723-280	Sequence 280, App
10	16.8	84.0	86574	18 US-10-775-169-173	Sequence 173, Appl
11	16.4	82.0	18	16 US-10-133-779-64	Sequence 64, Appl

12	16.4	82.0	18	17 US-10-271-602B-104	Sequence 104, App
13	16.4	82.0	515	18 US-10-767-701-19462	Sequence 19462, A
14	16.4	82.0	759	17 US-10-424-599-30744	Sequence 30744, A
15	16.4	82.0	1155	13 US-10-044-030-562	Sequence 562, App
16	16.4	82.0	2490	18 US-10-437-963-53682	Sequence 53682, A
17	16.4	82.0	2661	18 US-10-437-963-70761	Sequence 70761, A
18	16.4	82.0	3357	17 US-10-158-057-353	Sequence 353, App
19	16.4	82.0	135638	16 US-10-314-657-1	Sequence 1, Appl
20	16.4	82.0	135638	19 US-10-473-193-1	Sequence 1, Appl
21	16.4	82.0	439892	13 US-10-087-132-454	Sequence 454, App
22	15.8	79.0	196	18 US-10-425-115-111219	Sequence 111219, A
23	15.8	79.0	294	9 US-09-783-590-1155	Sequence 1155, Ap
24	15.8	79.0	474	13 US-10-027-632-281435	Sequence 281435, A
25	15.8	79.0	474	17 US-10-027-632-281435	Sequence 281435, A
26	15.8	79.0	476	18 US-10-425-115-11307	Sequence 11307, A
27	15.8	79.0	504	18 US-10-437-963-77414	Sequence 77414, A
28	15.8	79.0	523	17 US-10-231-956A-417	Sequence 417, App
29	15.8	79.0	618	13 US-10-027-632-134123	Sequence 134123, A
30	15.8	79.0	618	17 US-10-027-632-134123	Sequence 134123, A
31	15.8	79.0	621	13 US-10-027-632-106798	Sequence 106798, A
32	15.8	79.0	621	13 US-10-027-632-106798	Sequence 106798, A
33	15.8	79.0	621	17 US-10-027-632-134124	Sequence 134124, A
34	15.8	79.0	621	17 US-10-027-632-134124	Sequence 134124, A
35	15.8	79.0	724	13 US-10-027-632-24674	Sequence 24674, A
36	15.8	79.0	724	13 US-10-027-632-24674	Sequence 24674, A
37	15.8	79.0	724	17 US-10-027-632-24674	Sequence 24674, A
38	15.8	79.0	724	17 US-10-027-632-24674	Sequence 24674, A
39	15.8	79.0	773	18 US-10-425-115-102405	Sequence 102405, A
40	15.8	79.0	848	17 US-10-282-122A-37145	Sequence 37145, A
41	15.8	79.0	924	18 US-10-437-963-27553	Sequence 27553, A
42	15.8	79.0	987	17 US-10-282-122A-25867	Sequence 25867, A
43	15.8	79.0	1021	17 US-10-321-039-7	Sequence 7, Appl
44	15.8	79.0	1215	18 US-10-437-963-85599	Sequence 85599, A
45	15.8	79.0	1217	17 US-10-369-493-27595	Sequence 27595, A

ALIGNMENTS

RESULT 1
US-09-764-853-313
; Sequence 313, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 313
; LENGTH: 322
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (87)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (105)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (123)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-853-313

Query Match 84.0%; Score 16.8; DB 9; Length 322;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20

Db 278 AGCCCCAGCGCGCTCCAGCTG 297
| ||||| ||||| |||||

RESULT 2

US-10-432-934-17/c

; Sequence 17, Application US/10432934
; Publication No. US20040110269A1

; GENERAL INFORMATION:

; APPLICANT: Vipond, Richard

; APPLICANT: Shuttleworth, Helen

; APPLICANT: Ambrose, Emma

; APPLICANT: Minton, Nigel Peter

; TITLE OF INVENTION: Protection against mycobacterial infections

; FILE REFERENCE: 1581.097001

; CURRENT APPLICATION NUMBER: US/10/432,934

; CURRENT FILING DATE: 2003-05-28

; PRIOR APPLICATION NUMBER: PCT/GB01/05250

; PRIOR FILING DATE: 2001-11-28

; PRIOR APPLICATION NUMBER: GB0028966.0

; PRIOR FILING DATE: 2000-11-28

; NUMBER OF SEQ ID NOS: 118

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 17

; LENGTH: 1545

; TYPE: DNA

; ORGANISM: Mycobacterium tuberculosis

US-10-432-934-17

Query Match

84.0%; Score 16.8; DB 18; Length 1545;

Best Local Similarity 90.0%; Pred. No. 1.3e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20

||| ||||| ||||| |||||

Db 1236 ATCCAGCGCGCTCCAGCTG 1217

RESULT 3

US-10-027-632-97445/c

; Sequence 97445, Application US/10027632

; Publication No. US20020198371A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 97445

; LENGTH: 1782

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-97445

Query Match

84.0%; Score 16.8; DB 13; Length 1782;

Best Local Similarity 90.0%; Pred. No. 1.3e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20

||| ||||| ||||| |||||

Db 1781 AGCCCCAGCGCTCCAGCTG 1762

; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97445
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97445

Query Match 84.0%; Score 16.8; DB 17; Length 1782;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1781 AGCCAGCGCTCTCCAGCTG 1762

RESULT 6
US-10-027-632-97446/c
; Sequence 97446, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97446
; LENGTH: 1782
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97446

Query Match 84.0%; Score 16.8; DB 17; Length 1782;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 1781 AGCCAGCGCTCTCCAGCTG 1762

RESULT 7
US-09-764-853-77/c
; Sequence 77, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: Puz06
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 77
; LENGTH: 3957
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-853-77

Query Match 84.0%; Score 16.8; DB 9; Length 3957;
Best Local Similarity 90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 36 AGCCAGCGGCGCCAGCTG 17

RESULT 8
US-10-723-860-5186/c
; Sequence 5186, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: Methods for Screening for Soft Tissue Sarcoma Modulators
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5186
; LENGTH: 4309
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5186

Query Match 84.0%; Score 16.8; DB 18; Length 4309;
Best Local Similarity 90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 270 AGCCAGCGGCGCCAGCTG 251

RESULT 9
US-09-997-722-280/c
; Sequence 280, Application US/09997722
; Publication No. US20040072154A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71171/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/997,722
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 301
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 280
; LENGTH: 40304
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (14527)..(15049)
; OTHER INFORMATION: "n" at position 14527 through 15049 can be any base.
US-09-997-722-280


```
RESULT 14
US-10-424-599-30744/c
; Sequence 30744, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT FILING DATE: 2003-04-28
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 30744
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_127768C.1
US-10-424-599-30744

Query Match      82.0%; Score 16.4; DB 17; Length 759;
Best Local Similarity 94.4%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2  TCCAGCGCGCTCCAGCT 19
      ||||| ||||| ||||| |||||
DB      485  TCCATCGCGCTCCAGCT 468

RESULT 15
US-10-044-090-562/c
; Sequence 562, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 562
; LENGTH: 1155
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1383423.1
; NAME/KEY: unsure
; LOCATION: 344, 394
; OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-562

Query Match      82.0%; Score 16.4; DB 13; Length 1155;
Best Local Similarity 94.4%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2  TCCAGCGCGCTCCAGCT 19
      ||||| ||||| ||||| |||||
DB      632  TACCAGCGCGCTCCAGCT 615

RESULT 16
US-10-437-963-53682
; Sequence 53682, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ05C1
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 53682
; LENGTH: 2490
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_5585C.1
US-10-437-963-53682

Query Match      82.0%; Score 16.4; DB 18; Length 2490;
Best Local Similarity 94.4%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  CCCAGCGCGCTCCAGCTG 20
      ||||| ||||| ||||| |||||
DB      91  CCCCGCGCGCTCCAGCTG 108

RESULT 17
US-10-437-963-70761
; Sequence 70761, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 70761
; LENGTH: 2661
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71300C.1
US-10-437-963-70761

Query Match      82.0%; Score 16.4; DB 18; Length 2661;
Best Local Similarity 94.4%; Pred. No. 1.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3  CCCAGCGCGCTCCAGCTG 20
      ||||| ||||| ||||| |||||
DB      433  CCCCGCGCGCTCCAGCTG 450

RESULT 18
US-10-158-057-353/c.
; Sequence 353, Application US/10158057
; Publication No. US20040014039A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJZ05C1
```

; CURRENT APPLICATION NUMBER: US/10/158,057
; CURRENT FILING DATE: 2002-06-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 384
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 353
; LENGTH: 3357
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-158-057-353

Query Match 82.0%; Score 16.4; DB 17; Length 3357;
Best Local Similarity 94.4%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGGCTCCAGCT 19
| | | | | | | | | | | | | | | | | | | | |
Db 1059 TACCAGCGGCTCCAGCT 1042

RESULT 19
US-10-314-657-1
; Sequence 1, Application US/10314657
; Publication No. US20030175888A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, Ben
; APPLICANT: CHENG, Yi-Qiang
; APPLICANT: TANG, Gong-Li
; TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polyketide
; TITLE OF INVENTION: Synthases and Methods of Use
; FILE REFERENCE: 054030-0021
; CURRENT APPLICATION NUMBER: US/10/314,657
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: PCT/US02/08937
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US 60/278,935
; PRIOR FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 135638
; TYPE: DNA
; ORGANISM: Streptomyces atroolivaceus
US-10-314-657-1

Query Match 82.0%; Score 16.4; DB 16; Length 135638;
Best Local Similarity 94.4%; Pred. No. 98;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | |
Db 23908 CTCAGCGGCTCCAGCTG 23925

RESULT 20
US-10-473-193-1
; Sequence 1, Application US/10473193
; Publication No. US20050080247A1
; GENERAL INFORMATION:
; APPLICANT: SHEN, BEN
; APPLICANT: CHEN, Yi-Qiang
; APPLICANT: TANG, GONG-LI
; TITLE OF INVENTION: LEINAMYCIN BIOSYNTHESIS GENE CLUSTER AND ITS COMPONENTS AND THEIR
; TITLE OF INVENTION: USES
; FILE REFERENCE: 309T-000110US
; CURRENT APPLICATION NUMBER: US/10/473,193
; CURRENT FILING DATE: 2003-09-24
; PRIOR APPLICATION NUMBER: US 60/278,935
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/08937
; PRIOR FILING DATE: 2002-03-22
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1
; LENGTH: 135638
; TYPE: DNA
; ORGANISM: Streptomyces atroolivaceus
US-10-473-193-1

Query Match 82.0%; Score 16.4; DB 19; Length 135638;
Best Local Similarity 94.4%; Pred. No. 98;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | |
Db 23908 CTCAGCGGCTCCAGCTG 23925

RESULT 21
US-10-087-192-454
; Sequence 454, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 454
; LENGTH: 439892
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) -(439892)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-454

Query Match 82.0%; Score 16.4; DB 13; Length 439892;
Best Local Similarity 94.4%; Pred. No. 81;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCCAGCGGCTCCAGCTG 20
| | | | | | | | | | | | | | | | | | | | |
Db 371549 CCCAGCGGCTCCAGCTG 371566

RESULT 22
US-10-425-115-111219
; Sequence 111219, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 111219
; LENGTH: 196
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure

```
; LOCATION: (1)..(196)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRY4577_32922C.1
US-10-425-115-111219

Query Match          79.0%; Score 15.8; DB 18; Length 196;
Best Local Similarity 89.5%; Pred. No. 5.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCT 19
    |||||
Db 56 ATCCAGCGCGCTCAAGCT 74

RESULT 23
US-09-783-590-1155
; Sequence 1155, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: FO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1155
; LENGTH: 294
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (187)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (194)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (215)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (269)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-1155

Query Match          79.0%; Score 15.8; DB 9; Length 294;
Best Local Similarity 89.5%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCTG 20
    |||||
Db 164 TCCAGCTCCCTCCAGCTG 182

RESULT 24
US-10-027-632-281435
; Sequence 281435, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632

Query Match          79.0%; Score 15.8; DB 17; Length 474;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCTG 20
    |||||
Db 148 TCCAGCGGTCTCCATCTG 166
```

```
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281435
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281435

Query Match          79.0%; Score 15.8; DB 13; Length 474;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCTG 20
    |||||
Db 148 TCCAGCGGTCTCCATCTG 166

RESULT 25
US-10-027-632-281435
; Sequence 281435, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281435
; LENGTH: 474
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281435

Query Match          79.0%; Score 15.8; DB 17; Length 474;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGCTG 20
    |||||
Db 148 TCCAGCGGTCTCCATCTG 166
```

```
RESULT 26
US-10-425-115-11307/c
; Sequence 11307, Application US/10425115
; Publication No. US200400214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 11307
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_11030C.1
US-10-425-115-11307

Query Match          79.0%; Score 15.8; DB 18; Length 476;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TCCACGCGCGCTCCAGCTG 20
||||| ||||| ||||| ||||| |||||
Db 251 TCCCGCCGCGCTCCAGCTG 233

RESULT 27
US-10-437-963-77414
; Sequence 77414, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 77414
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77314C.1
US-10-437-963-77414

Query Match          79.0%; Score 15.8; DB 18; Length 504;
Best Local Similarity 89.5%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCACGCGCGCTCCAGCT 19
||||| ||||| ||||| ||||| |||||
Db 80 ACCCCAGCGCGCACCAGCT 98

RESULT 28
US-10-231-956A-417/c
; Sequence 417, Application US/10231956A
; Publication No. US20040053233A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Lorens, James B.
; APPLICANT: Xu, Weiduan
; APPLICANT: Bogenberger, Jakob
; APPLICANT: Holland, Sacha
; APPLICANT: Rigel Pharmaceuticals, Incorporated
; TITLE OF INVENTION: Modulators of Angiogenesis
; FILE REFERENCE: 021044-004100US
; CURRENT APPLICATION NUMBER: US/10/231,956A
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 417
; LENGTH: 523
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (1)...(523)
; OTHER INFORMATION: n = g, a, c or t
US-10-231-956A-417

Query Match          79.0%; Score 15.8; DB 17; Length 523;
Best Local Similarity 85.0%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCCACGCGCGCTCCAGCTG 20
||||| ||||| ||||| ||||| |||||
Db 20 ATCCGAGCTCGNTCCAGCTG 1

RESULT 29
US-10-027-632-134123
; Sequence 134123, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134123
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134123

Query Match          79.0%; Score 15.8; DB 13; Length 618;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCACGCGCGCTCCAGCT 19
||||| ||||| ||||| ||||| |||||
Db 419 ATCCCGCGCGCTCTCCAGCT 437
```



```
RESULT 30
US-10-027-632-134123
; Sequence 134123, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134123
; LENGTH: 618
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134123

Query Match 79.0%; Score 15.8; DB 17; Length 618;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCT 19
DB 419 ATCCCGCGGCTCTCCAGCT 437

RESULT 31
US-10-027-632-106798
; Sequence 106798, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106798
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106798

Query Match 79.0%; Score 15.8; DB 13; Length 621;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCT 19
DB 61 ATCCCGCGGCTCTCCAGCT 79

RESULT 32
US-10-027-632-134124
; Sequence 134124, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134124
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134124

Query Match 79.0%; Score 15.8; DB 13; Length 621;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCT 19
DB 61 ATCCCGCGGCTCTCCAGCT 79

RESULT 33
US-10-027-632-106798
; Sequence 106798, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
```

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106798
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134124

Query Match 79.0%; Score 15.8; DB 17; Length 621;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAGCT 19
||||| ||||| ||||| ||||| |||||
DB 61 ATCCCGCGCGCTCCAGCT 79

RESULT 34
US-10-027-632-134124
; Sequence 134124, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134124
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134124

Query Match 79.0%; Score 15.8; DB 17;
Best Local Similarity 89.5%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAGCT 19
||||| ||||| ||||| ||||| |||||
DB 61 ATCCCGCGCGCTCCAGCT 79

RESULT 35
US-10-027-632-24674/c
; Sequence 24674, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24674
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-24674

Query Match 79.0%; Score 15.8; DB 13; Length 724;
Best Local Similarity 89.5%; Pred. No. 4.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAGCT 19
||||| ||||| ||||| ||||| |||||
DB 562 ATCCCGCGCGCTCCAGCT 544

RESULT 36
US-10-027-632-24675/c
; Sequence 24675, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24675
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-24675

Query Match 79.0%; Score 15.8; DB 13;
Best Local Similarity 89.5%; Pred. No. 4.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAGCT 19
||||| ||||| ||||| ||||| |||||
DB 562 ATCCCGCGCGCTCCAGCT 544

```
RESULT 37
US-10-027-632-24674/c
; Sequence 24674, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24674
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-24674

Query Match          79.0%; Score 15.8; DB 17; Length 724;
Best Local Similarity 89.5%; Pred. No. 4.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGGCTCCAGCT 19
      ||||| ||||| ||||| |||||
Db      562 ATCCCGCGGCTCTCCAGCT 544

RESULT 38
US-10-027-632-24675/c
; Sequence 24675, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24675
; LENGTH: 724
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-24675

Query Match          79.0%; Score 15.8; DB 17; Length 724;
Best Local Similarity 89.5%; Pred. No. 4.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGGCTCCAGCT 19
      ||||| ||||| ||||| |||||
Db      562 ATCCCGCGGCTCTCCAGCT 544
```

```
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-24675

Query Match          79.0%; Score 15.8; DB 17; Length 724;
Best Local Similarity 89.5%; Pred. No. 4.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGGCTCCAGCT 19
      ||||| ||||| ||||| |||||
Db      562 ATCCCGCGGCTCTCCAGCT 544

RESULT 39
US-10-425-115-102405
; Sequence 102405, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; Plasmids
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 102405
; LENGTH: 773
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(773)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MPT4577_2489C.1
US-10-425-115-102405

Query Match          79.0%; Score 15.8; DB 18; Length 773;
Best Local Similarity 89.5%; Pred. No. 4.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TCCGAGCGGCTCCAGCTG 20
      ||||| ||||| ||||| |||||
Db      23 TGCCAGCGGCTCCGCGCTG 41

RESULT 40
US-10-282-122A-37145/c
; Sequence 37145, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
```

; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37145
; LENGTH: 848
; TYPE: DNA
; ORGANISM: Salmonella paratyphi A
US-10-282-122A-37145

Query Match 79.0%; Score 15.8; DB 17; Length 848;
Best Local Similarity 89.5%; Pred. NO. 4.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGCGCTCCAGCT 19
Db 685 ATACCAGCGCGCCAGCT 667

Search completed: June 4, 2005, 12:19:59
Job time : 192.651 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:32:00 ; Search time 1221.07 Seconds
(without alignments)
623.460 Million cell updates/sec

Title: US-09-674-277-23

Perfect score: 20

Sequence: 1 atccacgcgcgtccagctg 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	18	90.0	747	4	BG386940 602455130
C 2	17.4	87.0	644	5	BQ205315 UI-R-D21-
C 3	17.4	87.0	1825	9	CL490495 SAIL 539
C 4	17.4	87.0	3384	9	AY421230 Homo sapi
C 5	17.4	87.0	3384	9	AY421231 Pan trogl
C 6	17	85.0	312	2	BF231091 253318 BA
C 7	17	85.0	422	4	BI682647 463789 MA
C 8	17	85.0	496	2	BE482073 167730 BA
C 9	17	85.0	803	9	CG811669 FSAAU03TR
C 10	17	85.0	828	9	CG809419 FSAAU03TR
C 11	16.8	84.0	279	1	AA077608 7B27F10 C
C 12	16.8	84.0	365	1	AU186763 AU186763
C 13	16.8	84.0	368	1	AU189431 AU189431
C 14	16.8	84.0	379	6	CF066580 Ac449 Amp
C 15	16.8	84.0	420	5	BY003220 BY003220
C 16	16.8	84.0	447	1	AI147447 3961408.1
C 17	16.8	84.0	460	1	AU189676 AU189676
C 18	16.8	84.0	477	8	AZ261971 RPCI-23-1
C 19	16.8	84.0	485	1	AU187704 AU187704
C 20	16.8	84.0	487	1	AU188027 AU188027
C 21	16.8	84.0	490	1	AU186803 AU186803
C 22	16.8	84.0	496	1	AV435023 AV435023
C 23	16.8	84.0	497	2	BF738520 PM1-KT004
C 24	16.8	84.0	500	1	AU187012 AU187012

25	16.8	84.0	508	1	AU192619	AU192619
26	16.8	84.0	510	1	AU187727	AU187727
27	16.8	84.0	516	1	AV429796	AV429796
28	16.8	84.0	519	1	AU194260	AU194260
29	16.8	84.0	520	9	CR052332	Reverse s
30	16.8	84.0	521	1	AU190031	AU190031
C 31	16.8	84.0	521	8	AZ027551	RPCI-23-3
32	16.8	84.0	522	1	AV430388	AV430388
33	16.8	84.0	523	1	AU189534	AU189534
34	16.8	84.0	524	1	AU188665	AU188665
35	16.8	84.0	524	1	AV430179	AV430179
36	16.8	84.0	524	1	AV434390	AV434390
37	16.8	84.0	524	1	AV434834	AV434834
38	16.8	84.0	526	1	AU190978	AU190978
39	16.8	84.0	526	1	AV435285	AV435285
40	16.8	84.0	530	1	AV431724	AV431724
41	16.8	84.0	531	1	AV432654	AV432654
42	16.8	84.0	531	1	AV434483	AV434483
43	16.8	84.0	537	1	AV434259	AV434259
44	16.8	84.0	538	1	AV432097	AV432097
45	16.8	84.0	539	1	AV434239	AV434239

ALIGNMENTS

RESULT 1
BG386940/c
LOCUS
DEFINITION
602455130F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4583551 5',
mRNA sequence.
ACCESSION
BG386940
VERSION
BG386940.1 GI:13280295
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 747)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/
TITLES
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1307 row: k column: 08
High quality sequence stop: 692.

FEATURES
source

Location/Qualifiers
1..747
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4583551"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/note="Organ: colon; Vector: pOTB7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"

ORIGIN

Query Match 90.0%; Score 18; DB 4; Length 747;

Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCACGCGCGCTCCAGCTG 20
|||||
Db 738 CCACGCGCGCTCCAGCTG 721

RESULT 2
LOCUS BQ205315 644 bp mRNA linear EST 12-AUG-2004
DEFINITION UI-R-DZ1-cni-h-19-0-UI.s1 NCI_CGAP_DZ1 Rattus norvegicus cDNA clone
IMAGE:7346085 3', mRNA sequence.

ACCESSION BQ205315
VERSION BQ205315.1 GI:20421780
KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 644)
AUTHORS Ronaldo, M.P., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabsb@mail.nih.gov

Oligo-dt track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
IMAGE (<http://image.llnl.gov>)
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1..644
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="IMAGE:7346085"
/tissue_type="Chondrosarcoma"
/dev_stage="37 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DZ1"
/note="Organ: Spine; Vector: pTT73D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI;
UI-R-DZ1 is a normalized cDNA library containing the
following tissue(s): Swam Rat Chondrosarcoma. The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dt primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CATTTCTTGTGA. The Rat cartilaginous tumor tissue was
provided by Dr. Jeff Stevens at the University of Iowa.
TAG_LIB=UI-R-DZ1
TAG_SEQ=CATTTCTTGTGA"

ORIGIN

Query Match 87.0%; Score 17.4; DB 5; Length 644;
Best Local Similarity 94.7%; Pred. No. 2.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCCAGCGCGCTCCAGCT 19
|||||
Db 295 ATCCCAGCGCGCTCCAGCT 313

RESULT 3
LOCUS CL490495/c

DEFINITION SAIL 539 F01.v2 SAIL Collection Arabidopsis thaliana genomic clone
SAIL_539_F01.v2, genomic survey sequence.

ACCESSION CL490495
VERSION CL490495.1 GI:45973937
KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1825)
AUTHORS Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,
Dietrich, B., Ho, P., Backward, J., Ko, C., Clarke, J.D., Cotton, D.,
Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B.,
Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.

A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)
22356987
12468722
Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com

ABRC Stock Number CS822855; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.
Location/Qualifiers
1..1825
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL 539 F01 v2"
/clone_lib="SAIL Collection"
/note="T-DNA left border sequences were isolated using a
modified TAIL-PCR strategy"

FEATURES
source

Query Match 87.0%; Score 17.4; DB 9; Length 1825;
Best Local Similarity 94.7%; Pred. No. 2.8e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

ORIGIN

QY 1 ATCCCAGCGCGCTCCAGCT 19
|||||
Db 936 ATCCCAGCGCGCTCCAGCT 908

RESULT 4
LOCUS AY421230/c

DEFINITION Homo sapiens HCM7488 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.

ACCESSION AY421230
VERSION AY421230.1 GI:39777187
KEYWORDS GSS.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3384)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3384)
 AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence as made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 source
 Location/Qualifiers
 1..3384
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>3384
 /locus_tag="HCM7488"
 gene
 ORIGIN
 Query Match 87.0%; Score 17.4; DB 9; Length 3384;
 Best Local Similarity 94.7%; Pred. No. 2.8e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATCCCGAGCGCTCCAGCT 19
 |||||
 Db 101 ATCCCGAGCTCGTCCAGCT 83
 |||||
 RESULT 5
 AY421231/c
 LOCUS 3384 bp DNA linear GSS 12-DEC-2003
 DEFINITION Pan troglodytes HCM7488 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY421231
 VERSION AY421231.1 GI:39777188
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 REFERENCE 1 (bases 1 to 3384)
 AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 3384)
 AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence as made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 source
 Location/Qualifiers
 1..3384
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>3384
 gene

ORIGIN
 Query Match 87.0%; Score 17.4; DB 9; Length 3384;
 Best Local Similarity 94.7%; Pred. No. 2.8e+03;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATCCCGAGCGCTCCAGCT 19
 |||||
 Db 101 ATCCCGAGCTCGTCCAGCT 83
 |||||
 RESULT 6
 BF231091
 LOCUS 312 bp mRNA linear EST 27-MAR-2003
 DEFINITION 253318 BARC SBOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BF231091
 VERSION BF231091.1 GI:11169748
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 312)
 AUTHORS Sonstegard, T., Capuco, A.V., White, J., Van Tassel, C.P., Connor, E.E., Cho, J., Sultana, R., Shade, L., Wray, J.E., Wells, K.D. and Quackenbush, J.
 TITLE Analysis of bovine mammary gland EST and functional annotation of the Bos taurus gene index
 JOURNAL Mamm. Genome 13 (7), 373-379 (2002)
 MEDLINE 22135956
 PUBMED 12140684
 COMMENT Contact: Sonstegard TS
 USDA, ARS, Beltsville Agricultural Research Center
 Bldg. 200 Rm 2A, Beltsville, MD 20705, USA
 Tel: 301 504 8416
 Fax: 301 504 8414
 Email: tats@psi.barc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACACGCTATGACCAT
 BACKWARD: GTTTCGAGTCAGCAGC
 Plate: 107 row: F column: 7
 Seq primer: ATTAGTGACACTATAG.
 Location/Qualifiers
 1..312
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="BARC SBOV"
 /note="Vector: PCMV SPOR76; Site_1: NotI; Site_2: SalI;
 Library made from pooled mRNA isolated from mammary tissues at eight physiological, developmental, and disease states."
 ORIGIN
 Query Match 85.0%; Score 17; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 4.2e+03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 CCCAGCGCTCCAGCT 19
 |||||
 Db 92 CCCAGCGCTCCAGCT 108
 |||||
 RESULT 7
 BI682647
 LOCUS 422 bp mRNA linear EST 17-SEP-2001

```

DEFINITION 463789 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BI682647
VERSION BI682647.1 GI:15635581
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus

REFERENCE
AUTHORS Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keefe,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
PUBMED 11282978
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTCACGACG
Plate: 133 row: A column: 18
Seq primer: ATTAGTGCACATATAG.
Location/Qualifiers
1..422
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from lymph node, ovary, fat, hypothalamus, and pituitary."

FEATURES
source
QY 3 CCCAGCGCGCTCCAGCT 19
Db 29 CCCAGCGCGCTCCAGCT 45

ORIGIN
Query Match 85.0%; Score 17; DB 4; Length 422;
Best Local Similarity 100.0%; Pred. NO. 4.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCCAGCGCGCTCCAGCT 19
Db 29 CCCAGCGCGCTCCAGCT 45

ORIGIN
Query Match 85.0%; Score 17; DB 4; Length 422;
Best Local Similarity 100.0%; Pred. NO. 4.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCCAGCGCGCTCCAGCT 19
Db 29 CCCAGCGCGCTCCAGCT 45

ORIGIN
Query Match 85.0%; Score 17; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 4.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCCAGCGCGCTCCAGCT 19
Db 125 CCCAGCGCGCTCCAGCT 141

RESULT 9
CG811669/c
LOCUS CG811669
DEFINITION FSAU03TR LargeInsertGenomicLibrary Fusarium virguliforme genomic clone KMPv6A5, genomic survey sequence.
ACCESSION CG811669
VERSION CG811669.1 GI:38265143
KEYWORDS GSS.
SOURCE CG811669.1
ORGANISM Fusarium virguliforme
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.
REFERENCE 1 (bases 1 to 803)
AUTHORS Meksem,K., Ishihara,H., Koo,H., Shultz,J., Ali,S., Iqbal,J., Lightfoot,D.A. and Town,C.D.
TITLE End sequencing of BACs from a fingerprint physical map of the causative agent of soybean sudden death syndrome, Fusarium virguliforme
JOURNAL Unpublished (2003)
COMMENT Other GSSs: FSAU03TF
Contact: Chris Town and K. Meksem
The Center of Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale and Plant Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415, USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 3103 and 301-838-3523
Fax: 618 453-7457 and 301-838-0208
Email: meksemk@siu.edu; cdtown@tigr.org (URL: http://Fusariumvirguliform.siu.edu)
Seq primer: CAGGAACAGCTATGACC
Class: BAC ends.

```


FEATURES
source

Location/Qualifiers
1. .803
/organism="Fusarium virguliforme"
/mol_type="genomic DNA"
/cultivar="Monticello"
/db_xref="taxon:232082"
/clone="KMFv6A5"
/clone_lib="LargeInsertGenomicLibrary"
/note="Organ: Hyphae; Vector: pINDIGOBAC5; A single spore derived culture was used. Hyphae were grown in an incubator for four days. Nuclei were isolated and embedded in agarose, restriction digested with Hind III. Large size DNA fragments were ligated in vector pINDIGOBAC5 and electro-transformed into DH10B cells."

ORIGIN

Query Match 85.0%; Score 17; DB 9; Length 803;
Best Local Similarity 100.0%; Pred. No. 4.2e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCAGCGCGCTCCAGCT 19
|||||
Db 754 CCAGCGCGCTCCAGCT 738

RESULT 10

CG809419/c
LOCUS
DEFINITION
FSAB71TFB LargeInsertGenomicLibrary Fusarium virguliforme genomic
clone KMFv1K22, genomic survey sequence.

ACCESSION
CG809419
VERSION
GSS.
KEYWORDS
SOURCE
ORGANISM
Fusarium virguliforme

REFERENCE
AUTHORS
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.
1 (bases 1 to 828)

TITLE
Meksem, K., Ishihara, H., Koo, H., Shultz, J., Ali, S., Iqbal, J.,
Lightfoot, D.A. and Town, C.D.
End sequencing of BACs from a fingerprint physical map of the
causative agent of soybean sudden death syndrome, Fusarium
virguliforme

JOURNAL
COMMENT

Unpublished (2003)
Other_GSSs: FSAB71TRB
Contact: Chris Town and K. Meksem
The Center of Excellence in Soybean Research, Teaching and
Outreach, Southern Illinois University at Carbondale and Plant
Genomics, The Institute for Genomic Research
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 618 453 3103 and 301-838-3523
Fax: 618 453-7457 and 301-838-0208
Email: meksemk@siu.edu; cdtown@tigr.org (URL:
http://Fusariumvirguliforme.siu.edu)
Seq primer: CAGAAACAGCTATGACC
Class: BAC ends.

FEATURES
source

Location/Qualifiers
1. .828
/organism="Fusarium virguliforme"
/mol_type="genomic DNA"
/cultivar="Monticello"
/db_xref="taxon:232082"
/clone="KMFv1K22"
/clone_lib="LargeInsertGenomicLibrary"
/note="Organ: Hyphae; Vector: pINDIGOBAC5; A single spore derived culture was used. Hyphae were grown in an incubator for four days. Nuclei were isolated and embedded in agarose, restriction digested with Hind III. Large size DNA fragments were ligated in vector pINDIGOBAC5 and electro-transformed into DH10B cells."

ORIGIN

Query Match
Best Local Similarity
Matches

85.0%; Score 17; DB 9; Length 828;
100.0%; Pred. No. 4.2e+03;
17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CCAGCGCGCTCCAGCT 19
|||||
Db 783 CCAGCGCGCTCCAGCT 767

RESULT 11

AA077608/c
LOCUS
DEFINITION
AA077608
clone 7B27F10, mRNA sequence.

ACCESSION
AA077608
VERSION
EST.
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 279)

TITLE
Robbins, C.M., Nussbaum, J.C., Lovett, M. and Green, E.D.
2006 expressed-sequence tags derived from human chromosome
7-enriched cDNA libraries

JOURNAL
MEDLINE
PUBMED
COMMENT
Genome Res. 7 (3), 281-292 (1997)
97228905
9074931

Contact: Eric D. Green
Genome Technology Branch
National Human Genome Research Institute/NIH
49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892
Tel: 3014020201

Fax: 3014024735
Email: egreen@hgrl.nih.gov
Plate: 27 row: F column: 10
Seq primer: -21M13 (ABI)

FEATURES
source

Location/Qualifiers
1. .279
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="7B27F10"
/sex="female and male mixture"
/tissue_type="brain"
/dev_stage="pool of 9 week and 12 week"
/lab_host="E. coli strain DH5 alpha"
/clone_lib="Chromosome 7 Fetal Brain cDNA Library"
/note="Organ: brain; Vector: pAMP10; cDNA was generated from cytoplasmic RNA using a mixture of random DNA hexamers and oligo(dT). From this pool of cDNA, human chromosome 7-enriched cDNA was isolated by direct cDNA selection using chromosome 7 genomic DNA (cosmids). The resulting direct-selected cDNA was cloned into a plasmid vector using a non-directional uracil DNA glycosylase (UDG)-mediated cloning strategy."

ORIGIN

Query Match 84.0%; Score 16.8; DB 1; Length 279;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
|||||
Db 203 AGCCAGCGCGCGCCAGCTG 184

RESULT 12

AUI86763
LOCUS
DEFINITION
AUI86763
365 bp mRNA linear EST 14-OCT-2003
Porphyrha yezoensis TU-1 sporophytes Porphyrha yezoensis
cDNA clone PF003907_r 5', mRNA sequence.

```

ACCESSION      AU186763
VERSION        AU186763.1  GI:31919728
KEYWORDS       EST.
SOURCE         Porphyra yezoensis
ORGANISM       Porphyra yezoensis
               Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
               Porphyra.
REFERENCE      1 (bases 1 to 365)
AUTHORS        Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
               Tabata,S.
TITLE          COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
               OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
               FREQUENCY ANALYSIS
JOURNAL        J. Phycol. 39 (5), 923-930 (2003)
COMMENT        Contact: Erika Asamizu
               The First Laboratory for Plant Gene Research
               Kazusa DNA Research Institute
               Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
               Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES       1..365
               source
               /organism="Porphyra yezoensis"
               /mol_type="mRNA"
               /strain="TU-1"
               /db_xref="taxon:2788"
               /clone="PF003907_r"
               /dev_stage="sporophytes"
               /clone_lib="Porphyra yezoensis TU-1 sporophytes"
ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 365;
Best Local Similarity 90.0%; Pred. NO. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCAGCGCGCTCCAGCTG 20
      |||||||||
Db      169 ATCCAGCGCGCGCCAGCTG 188

RESULT 13
LOCUS      AU189431
DEFINITION AU189431 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
            cDNA clone PF039a09_r 5', mRNA sequence.
ACCESSION  AU189431
KEYWORDS
SOURCE      Porphyra yezoensis
ORGANISM     Porphyra yezoensis
               Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
               Porphyra.
REFERENCE    1 (bases 1 to 368)
AUTHORS      Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
               Tabata,S.
TITLE        COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
               OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
               FREQUENCY ANALYSIS
JOURNAL      J. Phycol. 39 (5), 923-930 (2003)
COMMENT      Contact: Erika Asamizu
               The First Laboratory for Plant Gene Research
               Kazusa DNA Research Institute
               Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
               Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES     1..368
               source
               /organism="Porphyra yezoensis"
               /mol_type="mRNA"
               /strain="TU-1"
               /db_xref="taxon:2788"
               /clone="PF039a09_r"
               /dev_stage="sporophytes"
               /clone_lib="Porphyra yezoensis TU-1 sporophytes"
ORIGIN

```

```

Query Match      84.0%; Score 16.8; DB 1; Length 368;
Best Local Similarity 90.0%; Pred. NO. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCAGCGCGCTCCAGCTG 20
      |||||||||
Db      315 ATCCAGCGCGCCAGCTG 334

RESULT 14
LOCUS      CF066580
DEFINITION Ac449 Amphidinium carterae Amphidinium carterae cDNA clone Ac449
            3', mRNA sequence.
ACCESSION  CF066580
VERSION     CF066580.1  GI:39544276
KEYWORDS    EST.
SOURCE      Amphidinium carterae
ORGANISM     Amphidinium carterae
               Eukaryota; Alveolata; Dinophyceae; Gymnodiniales; Gymnodiniaceae;
               Amphidinium.
REFERENCE    1 (bases 1 to 379)
AUTHORS      Bachvaroff,T.R., Concepcion,G.T., Rogers,C.R., Herman,E.M. and
               Delwiche,C.F.
TITLE        Dinoflagellate expressed sequence tag data indicate massive
               transfer of chloroplast genes to the nuclear genome
               Protist 155 (1), 65-78 (2004)
JOURNAL      Contact: Charles Delwiche
               University of Maryland, College Park
               H.J. Patterson Hall, College Park, MD 20742, USA
               Tel: 301-405-8300
               Fax: 301-314-9082
               Email: delwiche@umd.edu
               Plate: 5 row: A column: 9
               Seq primer: CTCGTCCGAATTCG.
               Location/Qualifiers
               1..379
               /organism="Amphidinium carterae"
               /mol_type="mRNA"
               /strain="CCMP 1314"
               /db_xref="taxon:2961"
               /clone="Ac449"
               /clone_lib="Amphidinium carterae"
               /note="Vector: modified pBluescript SK+; Site_1: EcoRI;
               Site_2: NotI"
ORIGIN
Query Match      84.0%; Score 16.8; DB 6; Length 379;
Best Local Similarity 90.0%; Pred. NO. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCAGCGCGCTCCAGCTG 20
      |||||||||
Db      22 ATCCAGCGCGCTCCAGCTG 41

RESULT 15
LOCUS      BY003220/c
DEFINITION BY003220 RIKEN full-length enriched, 10 days neonate olfactory
            brain Mus musculus cDNA clone E53011F01 5', mRNA sequence.
ACCESSION  BY003220
VERSION     BY003220.1  GI:26063469
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM     Mus musculus
               Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
               1 (bases 1 to 420)
REFERENCE    1 (bases 1 to 420)
AUTHORS      Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,
               Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,
               Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,

```

Schönbach, C., Gojōbori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusic, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T. T., Gariboldi, M., Gispi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedirski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L. M., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Ponting, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, K., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalon, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hironane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genetic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',
GAGAGAGAGAGCGCGCAACTCGAGTTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5',
GAGAGAGAGATCTCGAGTCTAATTAATTAATTCGCCCGCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBlue-script KS(+) after bulk excision from Lambda FLC I."

ORIGIN

```
Query Match      84.0%; Score 16.8; DB 5; Length 420;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 ATCCAGCGGCTCCAGCTG 20
Dy 363 ACCCAGCGCTCTCCAGCTG 344

PRCIT.T 16

REF ID: A1147447/C

LOCUS

DEFINITION

1
2
3
4
5
6
7
8

ACCESSION

VERSION
REVISED

KEYWORDS
SOURCE

BOOKS
ORGANISM

53

REFERENCE

AUTHORS

ETITL

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

FEATURES

source

```

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 447;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
    ||||| ||||| ||||| |||||
Db 294 ATCCCACTGCGCTCCAGCTG 275

RESULT 17
AUI89676
LOCUS
DEFINITION
AUI89676 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
CDNA clone PF042c02_r 5', mRNA sequence.
AUI89676 460 bp mRNA linear EST 14-OCT-2003
AUI89676 1 GI:31925551
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Porphyra yezoensis
Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
1 (bases 1 to 460)
Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
Tabata,S.
COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
FREQUENCY ANALYSIS
J. Phycol. 39 (5), 923-930 (2003)
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
Location/Qualifiers
1..460
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PF042c02_r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 460;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
    ||||| ||||| ||||| |||||
Db 129 ATCCGAGCGGCCCGAGCTG 148

RESULT 18
AZ261971/c
LOCUS
DEFINITION
RPCI-23-123P23-TV RPCI-23 Mus musculus genomic clone
RPCI-23-123P23, genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
GSS.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 477)
Zhao,S., Nierman,W., Feidblyum,T., Malek,J., Shatsman,S.,
Akinret,B., Levins,M., McGann,S., Teegaye,G., Geer,K., Krol,M., de
Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)

Other GSSs: RPCI-23-123P23.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac ends/mouse/bac\_end\_intro.html
Plate: 123 row: P column: 23
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..477
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-123P23"
/sex="Female"
/lab host="DH10B"
/clone_lib="RPCI-23"
/organism="Kidney/Brain; Vector: pBACe3.6; Site:1:
EcORI; Site 2: EcORI; Female CS7BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcORI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN
Query Match      84.0%; Score 16.8; DB 8; Length 477;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
    ||||| ||||| ||||| |||||
Db 356 ATCCCACTGCGCTCCAGCTG 337

RESULT 19
AUI87704
LOCUS
DEFINITION
AUI87704 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
CDNA clone PF016c10_r 5', mRNA sequence.
AUI87704 485 bp mRNA linear EST 14-OCT-2003
AUI87704 1 GI:31921601
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Porphyra yezoensis
Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
1 (bases 1 to 485)
Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
Tabata,S.
COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
FREQUENCY ANALYSIS
J. Phycol. 39 (5), 923-930 (2003)
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
Location/Qualifiers
1..485
/organism="Porphyra yezoensis"
/mol_type="mRNA"

```

```

/strain="TU-1"
/db_xref=taxon:2788"
/clone="PF016c10_r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN

Query Match      84.0%; Score 16.8; DB 1; Length 485;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGGCTCCAGCTG 20
        ||| ||||| ||||| |||||
DB      154 ATCCGAGCGGCCCGAGCTG 173

RESULT 20
AU188027
LOCUS      AU188027      487 bp      mRNA      linear      EST 14-OCT-2003
DEFINITION AU188027 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
            cDNA clone PF020e09_r 5', mRNA sequence.
ACCESSION  AU188027
VERSION     AU188027.1 GI:31922251
KEYWORDS   EST.
SOURCE     Porphyra yezoensis
ORGANISM   Porphyra yezoensis
            Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
            Porphyra.
REFERENCE  1 (bases 1 to 487)
AUTHORS   Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
            Tabata,S.
TITLE     COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
            OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
            FREQUENCY ANALYSIS
JOURNAL    J. Phycol. 39 (5), 923-930 (2003)
COMMENT    Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazuea DNA Research Institute
            Yana 1533-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazuea.or.jp, URL:http://www.kazuea.or.jp/en/plant/

FEATURES             Location/Qualifiers
     source            1..487
     organism="Porphyra yezoensis"
     mol_type="mRNA"
     strain="TU-1"
     db_xref="taxon:2788"
     clone="PF020e09_r"
     dev_stage="sporophytes"
     clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN

Query Match      84.0%; Score 16.8; DB 1; Length 487;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGGCTCCAGCTG 20
        ||| ||||| ||||| |||||
DB      174 ATCCGAGCGGCCCGAGCTG 193

RESULT 21
AU186803
LOCUS      AU186803      490 bp      mRNA      linear      EST 14-OCT-2003
DEFINITION AU186803 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
            cDNA clone PF004c06_r 5', mRNA sequence.
ACCESSION  AU186803
VERSION     AU186803.1 GI:31919807
KEYWORDS   EST.
SOURCE     Porphyra yezoensis
ORGANISM   Porphyra yezoensis
            Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
            Porphyra.
REFERENCE  1 (bases 1 to 490)

```

```

Db      301 ATGCCAGCGCGCCCGCTG 320
|||||||||
RESULT 23
BF738520
LOCUS      497 bp      mRNA      linear      EST 10-JAN-2001
DEFINITION PM1-KT0042-151200-003-e10 KT0042 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF738520
VERSION     BF738520.1 GI:12065196
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 497)
AUTHORS   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
          Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
          Goldmann,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          10737800
          CONTACT: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM1&t2=PM1-KT0042-
          151200-003-e10&t3=2000-12-15&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 36
          High quality sequence stop: 497.
          Location/Qualifiers
            1..497
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /dev_stage="Adult"
              /clone_lib="KT0042"
              /note="Organ: bladder tumor; Vector: puc18; Site 1: SmaI;
              Site 2: SmaI; A mini-library was made by cloning products
              derived from ORESTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the puc 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."
          ORIGIN
            Query Match      84.0%; Score 16.8; DB 2; Length 497;
            Best Local Similarity 90.0%; Pred. No. 5.2e+03;
            Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

            Qy      1 ATCCAGCGCGCTCCAGCTG 20
              |||||
            Db      149 AGCCCGCGCGCCCGCTG 168
              |||||

          RESULT 24
          AUI87012
          LOCUS      500 bp      mRNA      linear      EST 14-OCT-2003
          DEFINITION AUI87012 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
          cDNA clone PF007a12_r 5', mRNA sequence.
          ACCESSION  AUI87012
          KEYWORDS   EST.
          SOURCE     Porphyra yezoensis
          ORGANISM   Porphyra yezoensis
          Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
          Porphyra.
          1 (bases 1 to 508)
          Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
          Tabata,S.
          COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
          OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
          FREQUENCY ANALYSIS
          J. Phycol. 39 (5), 923-930 (2003)
          Contact: Erika Asamizu
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
          Location/Qualifiers
            1..508
              /organism="Porphyra yezoensis"
              /mol_type="mRNA"
              /strain="TU-1"
              /db_xref="taxon:2788"
              /clone="PF007a12_r"
              /dev_stage="sporophytes"
              /clone_lib="Porphyra yezoensis TU-1 sporophytes"
          ORIGIN
            Query Match      84.0%; Score 16.8; DB 1; Length 500;
            Best Local Similarity 90.0%; Pred. No. 5.2e+03;
            Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

            Qy      1 ATCCAGCGCGCTCCAGCTG 20
              |||||
            Db      404 ATGCCAGCGCGCCCGCTG 423
              |||||

          RESULT 25
          AUI92619
          LOCUS      508 bp      mRNA      linear      EST 14-OCT-2003
          DEFINITION AUI92619 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
          cDNA clone PFL033e05_r 5', mRNA sequence.
          ACCESSION  AUI92619
          VERSION     AUI92619.1 GI:31931440
          KEYWORDS   EST.
          SOURCE     Porphyra yezoensis
          ORGANISM   Porphyra yezoensis
          Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
          Porphyra.
          1 (bases 1 to 508)
          Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
          Tabata,S.
          COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
          OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
          FREQUENCY ANALYSIS
          J. Phycol. 39 (5), 923-930 (2003)
          Contact: Erika Asamizu
          The First Laboratory for Plant Gene Research
          Kazusa DNA Research Institute
          Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
          Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
          Location/Qualifiers
            1..508
              /organism="Porphyra yezoensis"
              /mol_type="mRNA"
              /strain="TU-1"
              /db_xref="taxon:2788"
              /clone="PFL033e05_r"
              /dev_stage="sporophytes"
              /clone_lib="Porphyra yezoensis TU-1 sporophytes"
          ORIGIN

```

Query Match 84.0%; Score 16.8; DB 1; Length 508;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
||| ||||| ||||| |||||
Db 401 ATGCCAGCGCGCCCGAGCTG 420

RESULT 26
AUI87727
LOCUS
DEFINITION AUI87727 Porphyra yezoensis 510 bp mRNA linear EST 14-OCT-2003
CDNA clone PF016e12_r 5', mRNA sequence.

ACCESSION AUI87727.1 GI:31921647
VERSION
KEYWORDS
SOURCE
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.

REFERENCE 1 (bases 1 to 510)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..510
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PF016e12_r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 510;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
||| ||||| ||||| |||||
Db 186 ATGCCAGCGCGCCCGAGCTG 205

RESULT 27
AV429796
LOCUS
DEFINITION AV429796 Porphyra yezoensis 516 bp mRNA linear EST 23-AUG-2000
FL008f03_r 5', mRNA sequence.

ACCESSION AV429796.1 GI:8585021
VERSION
KEYWORDS
SOURCE
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.

REFERENCE 1 (bases 1 to 516)
AUTHORS Nikaïdo,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and Tabata,S.
TITLE Generation of 10,154 expressed sequence tags from a leafy gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL DNA Res. 7, 223-227 (2000)
MEDLINE 20363100
PubMed 10907854

COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..516
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PL008f03_r"
/clone_lib="Porphyra yezoensis TU-1"
/note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 516;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
||| ||||| ||||| |||||
Db 475 ATGCCAGCGCGCCCGAGCTG 494

RESULT 28
AUI94260
LOCUS
DEFINITION AUI94260 Porphyra yezoensis 519 bp mRNA linear EST 14-OCT-2003
CDNA clone PFL057e04_r 5', mRNA sequence.

ACCESSION AUI94260
VERSION
KEYWORDS
SOURCE
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.

REFERENCE 1 (bases 1 to 519)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1..519
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL057e04_r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 519;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
||| ||||| ||||| |||||
Db 439 ATGCCAGCGCGCCCGAGCTG 458

RESULT 29
CR052332
LOCUS
CR052332 520 bp DNA linear GSS 05-JUL-2004

DEFINITION Reverse strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN20j18, genomic survey sequence.

ACCESSION CR052332
 VERSION 1
 KEYWORDS GSS; genome survey sequence; MICR.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 520)
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.
 TITLE Direct Submission
 JOURNAL Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. <http://www.sanger.ac.uk/MICR>

FEATURES
 source
 1..520
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHPN20j18"
 /clone_lib="MHPN"

ORIGIN

Query Match 84.0%; Score 16.8; DB 9; Length 520;
 Best Local Similarity 90.0%; Pred. No. 5.2e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGCGCTCCAGCTG 20

Db 439 ACCCGAGCGCGCTCCAGCTG 458

RESULT 30

AU190031

LOCUS AU190031 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
 DEFINITION cDNA clone PF046h02_r 5', mRNA sequence.

ACCESSION AU190031
 VERSION AU190031
 KEYWORDS EST.
 SOURCE Porphyra yezoensis
 ORGANISM Porphyra yezoensis

REFERENCE 1 (bases 1 to 521)
 AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and Tabata,S.
 TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG FREQUENCY ANALYSIS

J. Phycol. 39 (5), 923-930 (2003)
 CONTACT: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>.

FEATURES
 source

1..521
 Location/Qualifiers
 /organism="Porphyra yezoensis"
 /mol_type="mRNA"
 /strain="TU-1"
 /db_xref="taxon:2788"
 /clone="PF046h02_r"
 /dev_stage="sporophytes"
 /clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN

Query Match 84.0%; Score 16.8; DB 1; Length 521;
 Best Local Similarity 90.0%; Pred. No. 5.2e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGCGCTCCAGCTG 20

Db 57 ATCCGAGCGCGCGCCAGCTG 76

RESULT 31

AZ027551/c

LOCUS AZ027551/c

DEFINITION

RP01-23-343M24, genomic survey sequence.

ACCESSION AZ027551

VERSION AZ027551.1

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 521)

REFERENCE 1 (bases 1 to 521)

AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.

TITLE Mouse BAC End Sequences from Library RPCI-23

JOURNAL Unpublished (1999)

COMMENT Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)

or from Resea ch Genetics (info@resgen.com). BAC end page:

http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html

Plate: 343 row: M column: 24

Seq primer: SP6

Class: BAC ends.

FEATURES

source

1..521

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-343M24"

/sex="Female"

/lab_host="DH10B"

/clone_lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:

EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcoRI and EcoRI Methylase. Size

selected DNA was cloned into the pBACe3.6 vector at the

EcoRI sites. The ligation products were transformed into

DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 84.0%; Score 16.8; DB 8; Length 521;

Best Local Similarity 90.0%; Pred. No. 5.2e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGCGCTCCAGCTG 20

Db 360 ATCCGAGCGCGCTCCAGCTG 341

RESULT 32

AV430388

LOCUS AV430388

DEFINITION

Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone

PL017H04_r 5', mRNA sequence.

ACCESSION AV430388

VERSION AV430388.1 GI:8595613
KEYWORDS EST.
SOURCE Porphyra yezoensis
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 522)
AUTHORS Nikaïdo,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
Tabata,S.
TITLE Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL DNA Res. 7, 223-227 (2000)
MEDLINE 20363100
PUBMED 10907854
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..522
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PF017h04_r"
/clone_lib="Porphyra yezoensis TU-1"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 522;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCAGCGCGCTCCAGCTG 20
||| ||||| ||||| ||||| |||||
Db 437 ATCCAGCGCGCGCCAGCTG 456
||| ||||| ||||| ||||| |||||
RESULT 33
AUI89534
LOCUS
DEFINITION AUI89534 Porphyra yezoensis 523 bp mRNA linear EST 14-OCT-2003
CDNA clone PF040d06_r 5', mRNA sequence.
ACCESSION AUI89534
VERSION AUI89534.1 GI:31925264
KEYWORDS EST.
SOURCE Porphyra yezoensis
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 523)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..523
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PF040d06_r"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"
/dev_stage="sporophytes"

ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 523;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCAGCGCGCTCCAGCTG 20
||| ||||| ||||| ||||| |||||
Db 480 ATCCAGCGCGCGCCAGCTG 499
||| ||||| ||||| ||||| |||||
RESULT 34
AUI88665
LOCUS
DEFINITION AUI88665 Porphyra yezoensis 524 bp mRNA linear EST 14-OCT-2003
CDNA clone PF028h09_r 5', mRNA sequence.
ACCESSION AUI88665
VERSION AUI88665.1 GI:31923532
KEYWORDS EST.
SOURCE Porphyra yezoensis
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 524)
AUTHORS Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
Tabata,S.
TITLE COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
FREQUENCY ANALYSIS
JOURNAL J. Phycol. 39 (5), 923-930 (2003)
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES
source
1..524
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PF028h09_r"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"
/dev_stage="sporophytes"
ORIGIN
Query Match 84.0%; Score 16.8; DB 1; Length 524;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATCCAGCGCGCTCCAGCTG 20
||| ||||| ||||| ||||| |||||
Db 444 ATCCAGCGCGCGCCAGCTG 463
||| ||||| ||||| ||||| |||||
RESULT 35
AV430179
LOCUS
DEFINITION AV430179 Porphyra yezoensis 524 bp mRNA linear EST 23-AUG-2000
PL014f07_r 5', mRNA sequence.
ACCESSION AV430179
VERSION AV430179.1 GI:8585404
KEYWORDS EST.
SOURCE Porphyra yezoensis
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 524)
AUTHORS Nikaïdo,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
Tabata,S.
TITLE Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL DNA Res. 7, 223-227 (2000)

```

MEDLINE      20363100
PUBMED       10907854
COMMENT      Contact: Erika Asamizu
             The First Laboratory for Plant Gene Research
             Kazusa DNA Research Institute
             Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
             Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. .524
   /organism="Porphyra yezoensis"
   /mol_type="mRNA"
   /strain="TU-1"
   /db_xref="taxon:2788"
   /clone="PFO14f07 r"
   /clone_lib="Porphyra yezoensis TU-1"
   /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
   XhoI"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 524;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGCGCTCCAGCTG 20
      || ||||| ||||| |||||
Db      481 ATGCCAGCGCGCCCCAGCTG 500

RESULT 36
LOCUS    AV434390
DEFINITION AV434390 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PM042e08 r 5', mRNA sequence.
ACCESSION AV434390
VERSION   AV434390.1 GI:8589615
KEYWORDS EST.
SOURCE   Porphyra yezoensis
ORGANISM Porphyra yezoensis
          Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
          Porphyra.
REFERENCE 1 (bases 1 to 524)
AUTHORS  Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
          Tabata,S.
          Generation of 10,154 expressed sequence tags from a leafy
          gametophyte of a marine red alga, Porphyra yezoensis
          DNA Res. 7, 223-227 (2000)
CONTACT: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. .524
   /organism="Porphyra yezoensis"
   /mol_type="mRNA"
   /strain="TU-1"
   /db_xref="taxon:2788"
   /clone="PM042e08 r"
   /clone_lib="Porphyra yezoensis TU-1"
   /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
   XhoI"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 524;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGCGCTCCAGCTG 20
      || ||||| ||||| |||||
Db      497 ATGCCAGCGCGCCCCAGCTG 516

MEDLINE      20363100
PUBMED       10907854
COMMENT      Contact: Erika Asamizu
             The First Laboratory for Plant Gene Research
             Kazusa DNA Research Institute
             Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
             Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. .524
   /organism="Porphyra yezoensis"
   /mol_type="mRNA"
   /strain="TU-1"
   /db_xref="taxon:2788"
   /clone="PM042e08 r"
   /clone_lib="Porphyra yezoensis TU-1"
   /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
   XhoI"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 524;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGCGCTCCAGCTG 20
      || ||||| ||||| |||||
Db      497 ATGCCAGCGCGCCCCAGCTG 516
```

```

RESULT 37
LOCUS    AV434834
DEFINITION AV434834 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PM048h02_r 5', mRNA sequence.
ACCESSION AV434834
VERSION   AV434834.1 GI:8590059
KEYWORDS EST.
SOURCE   Porphyra yezoensis
ORGANISM Porphyra yezoensis
          Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
          Porphyra.
REFERENCE 1 (bases 1 to 524)
AUTHORS  Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
          Tabata,S.
          Generation of 10,154 expressed sequence tags from a leafy
          gametophyte of a marine red alga, Porphyra yezoensis
          DNA Res. 7, 223-227 (2000)
CONTACT: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. .524
   /organism="Porphyra yezoensis"
   /mol_type="mRNA"
   /strain="TU-1"
   /db_xref="taxon:2788"
   /clone="PM048h02 r"
   /clone_lib="Porphyra yezoensis TU-1"
   /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
   XhoI"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 524;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCGAGCGCGCTCCAGCTG 20
      || ||||| ||||| |||||
Db      494 ATGCCAGCGCGCCCCAGCTG 513

RESULT 38
LOCUS    AU190978
DEFINITION AU190978 Porphyra yezoensis TU-1 sporophytes Porphyra yezoensis
cDNA clone PFL009h02_r 5', mRNA sequence.
ACCESSION AU190978
VERSION   AU190978.1 GI:31928164
KEYWORDS EST.
SOURCE   Porphyra yezoensis
ORGANISM Porphyra yezoensis
          Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
          Porphyra.
REFERENCE 1 (bases 1 to 526)
AUTHORS  Asamizu,E., Nakajima,M., Kitade,Y., Saga,N., Nakamura,Y. and
          Tabata,S.
          COMPARISON OF RNA EXPRESSION PROFILES BETWEEN THE TWO GENERATIONS
          OF PORPHYRA YEZOENSIS (RHODOPHYTA), BASED ON EXPRESSED SEQUENCE TAG
          FREQUENCY ANALYSIS
          J. Phycol. 39 (5), 923-930 (2003)
CONTACT: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.

FEATURES
source
1. .526
```

```
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PFL009h02.r"
/dev_stage="sporophytes"
/clone_lib="Porphyra yezoensis TU-1 sporophytes"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 526;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
   |||||||
Db 441 ATGCCAGCGGCCCGAGCTG 460

RESULT 39
AV435285
LOCUS
DEFINITION AV435285 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PM055d01 r 5', mRNA sequence.
ACCESSION AV435285
VERSION AV435285.1 GI:8590510
KEYWORDS EST.
SOURCE Porphyra yezoensis
ORGANISM Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales;
Porphyra.
REFERENCE 1 (bases 1 to 526)
AUTHORS Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
Tabata,S.
TITLE Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL DNA Res. 7, 223-227 (2000)
MEDLINE 20363100
PUBMED 10907854
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..526
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PM055d01.r"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 526;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
   |||||||
Db 171 ATGCCAGCGGCCCGAGCTG 190

RESULT 40
AV431724
LOCUS
DEFINITION AV431724 Porphyra yezoensis TU-1 Porphyra yezoensis cDNA clone
PM004f09 r 5', mRNA sequence.
ACCESSION AV431724
VERSION AV431724.1 GI:8586949
KEYWORDS EST.
SOURCE Porphyra yezoensis
Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae;
Porphyra.
REFERENCE 1 (bases 1 to 530)
AUTHORS Nikaido,I., Asamizu,E., Nakajima,M., Nakamura,Y., Saga,N. and
Tabata,S.
TITLE Generation of 10,154 expressed sequence tags from a leafy
gametophyte of a marine red alga, Porphyra yezoensis
JOURNAL DNA Res. 7, 223-227 (2000)
MEDLINE 20363100
PUBMED 10907854
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..530
/organism="Porphyra yezoensis"
/mol_type="mRNA"
/strain="TU-1"
/db_xref="taxon:2788"
/clone="PM004f09.r"
/clone_lib="Porphyra yezoensis TU-1"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

ORIGIN
Query Match      84.0%; Score 16.8; DB 1; Length 530;
Best Local Similarity 90.0%; Pred. No. 5.2e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCGAGCGGCTCCAGCTG 20
   |||||||
Db 252 ATGCCAGCGGCCCGAGCTG 271

Search completed: June 4, 2005, 11:46:40
Job time : 1223.07 secs
```

This Page Blank (uspro)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:48:59 ; Search time 44.497 Seconds
(without alignments)
735.454 Million cell updates/sec

Title: US-09-674-277-23

Perfect score: 20

Sequence: 1 atccacagcgctccagctg 20

Scoring table: IDENTITY NJC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/prodata/1/ina/PCITUS_COMB.seq:*

6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17	85.0	1446	4	US-09-674-277-23
C 2	17	85.0	34199	4	US-09-320-878-19
C 3	16.8	84.0	601	4	US-09-320-878-19
C 4	16.8	84.0	4251	4	US-09-320-878-19
C 5	16.8	84.0	54484	4	US-09-320-878-19
C 6	16.8	84.0	101356	4	US-09-320-878-19
C 7	16.8	84.0	101357	4	US-09-320-878-19
C 8	16.8	84.0	4403765	3	US-09-103-840A-2
C 9	16.8	84.0	4411529	3	US-09-103-840A-2
C 10	16.4	82.0	18	4	US-09-747-391-64
C 11	15.8	79.0	424	4	US-09-270-767-10670
C 12	15.8	79.0	1260	4	US-09-320-878-19
C 13	15.8	79.0	2180	4	US-09-270-767-10685
C 14	15.8	79.0	2457	4	US-10-303-812-1
C 15	15.8	79.0	9035	4	US-09-320-878-19
C 16	15.8	79.0	14871	4	US-09-320-878-19
C 17	15.8	79.0	16157	4	US-09-320-878-19
C 18	15.4	77.0	315	4	US-09-252-991A-12396
C 19	15.4	77.0	1440	4	US-09-252-991A-12120
C 20	15.4	77.0	1494	4	US-09-320-878-19
C 21	15.4	77.0	1632	4	US-09-252-991A-12304
C 22	15.4	77.0	3177	4	US-09-320-878-19
C 23	15.4	77.0	6380	4	US-09-320-878-19
C 24	15.4	77.0	9551	1	US-08-056-200-93
C 25	15.4	77.0	9551	2	US-08-800-644-93
C 26	15.4	77.0	9880	4	US-09-320-878-19
C 27	15.4	77.0	13842	3	US-09-105-537-5

C 28	15.4	77.0	36778	3	US-09-105-537-5
C 29	15.4	77.0	38506	3	US-09-320-878-19
C 30	15.4	77.0	38506	4	US-09-141-908-1
C 31	15.4	77.0	38506	4	US-09-657-440-19
C 32	15.4	77.0	63183	4	US-09-949-016-13047
C 33	15.4	77.0	63183	4	US-09-949-016-13048
C 34	15.2	76.0	463	4	US-09-513-999C-4006
C 35	15.2	76.0	531	4	US-09-351-150A-40
C 36	15.2	76.0	601	4	US-09-949-016-39156
C 37	15.2	76.0	601	4	US-09-949-016-112208
C 38	15.2	76.0	601	4	US-09-949-016-112249
C 39	15.2	76.0	601	4	US-09-949-016-187754
C 40	15.2	76.0	601	4	US-09-949-016-187795
C 41	15.2	76.0	723	4	US-09-949-016-5895
C 42	15.2	76.0	1176	4	US-09-489-039A-5553
C 43	15.2	76.0	1257	4	US-09-489-039A-2959
C 44	15.2	76.0	1269	4	US-09-902-540-7168
C 45	15.2	76.0	1488	4	US-09-489-039A-3357

ALIGNMENTS

RESULT 1

US-09-902-540-5188/c

; Sequence 5188, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 5188

; LENGTH: 1446

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-5188

Query Match 85.0%; Score 17; DB 4; Length 1446;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGC 18

DB 420 TCCAGCGCGCTCCAGC 404

RESULT 2

US-09-902-540-1255

; Sequence 1255, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 1255

; LENGTH: 34199

; TYPE: DNA

```
; ORGANISM: Myxococcus xanthus
US-09-902-540-1255

Query Match      85.0%; Score 17; DB 4; Length 34199;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TCCACGGCGCTCCAGC 18
    |||||
Db 18253 TCCACGGCGCTCCAGC 18269

RESULT 3
US-09-949-016-179976/c
; Sequence 179976, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 179976
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-179976

Query Match      84.0%; Score 16.8; DB 4; Length 601;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
    |||||
Db 271 AGCCAGCGGTCTCCAGCTG 252

RESULT 4
US-09-902-540-5736
; Sequence 5736, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 5736
; LENGTH: 4251
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-5736

Query Match      84.0%; Score 16.8; DB 4; Length 4251;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
    |||||
Db 53980 AGCCAGCGGTCTCCAGCTG 53999

RESULT 5
US-09-902-540-1272
; Sequence 1272, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1272
; LENGTH: 54484
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1272

Query Match      84.0%; Score 16.8; DB 4; Length 54484;
Best Local Similarity 90.0%; Pred. No. 1.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
    |||||
Db 41319 ATCCAGCGGTCTCCAGCTG 41338

RESULT 6
US-09-949-016-12364
; Sequence 12364, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12364
; LENGTH: 101356
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12364

Query Match      84.0%; Score 16.8; DB 4; Length 101356;
Best Local Similarity 90.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATCCAGCGGCTCCAGCTG 20
    |||||
Db 53980 AGCCAGCGGTCTCCAGCTG 53999

RESULT 7
US-09-949-016-16924
; Sequence 16924, Application US/09949016
; Patent No. 6812339
```

```
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16924
; LENGTH: 101357
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16924

Query Match      84.0%; Score 16.8; DB 4; Length 101357;
Best Local Similarity 90.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATCCGAGCGGCTCCAGCTG 20
Db      53980 AGCCGAGCGGTGCTCCAGCTG 53999

RESULT 8
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match      84.0%; Score 16.8; DB 3; Length 4403765;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATCCGAGCGGCTCCAGCTG 20
Db      206631 ATCCGAGCGGCTCCAGGTG 206612

RESULT 9
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
```

```
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match      84.0%; Score 16.8; DB 3; Length 4411529;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 ATCCGAGCGGCTCCAGCTG 20
Db      206464 ATCGCAGCGGCTCCAGGTG 206445

RESULT 10
US-09-747-391-64
; Sequence 64, Application US/09747391
; Patent No. 6670124
; GENERAL INFORMATION:
; APPLICANT: Chow, Robert
; APPLICANT: Tonal, Richard
; APPLICANT: StemCyt, Inc.
; TITLE OF INVENTION: High Throughput Methods of HLA Typing
; FILE REFERENCE: 020035-000210US
; CURRENT APPLICATION NUMBER: US/09/747,391
; CURRENT FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/172,768
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 278
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-747-391-64

Query Match      82.0%; Score 16.4; DB 4; Length 18;
Best Local Similarity 94.4%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 TCCGAGCGGCTCCAGCT 19
Db      1 TACCAGCGGCTCCAGCT 18

RESULT 11
US-09-270-767-10670/c
; Sequence 10670, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10670
; LENGTH: 424
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; US-09-270-767-10670
```

```
Query Match          79.0%; Score 15.8; DB 4; Length 424;
Best Local Similarity 89.5%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TCCAGCGCGCTCCAGCTG 20
Db      156 TCGTAGCGCGCTCCAGCTG 138

RESULT 12
US-09-540-9533/c
; Sequence 9533, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9533
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-540-9533

Query Match          79.0%; Score 15.8; DB 4; Length 1260;
Best Local Similarity 89.5%; Pred. No. 3.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TCCAGCGCGCTCCAGCTG 20
Db      1147 TCCCGCGCGCTCCAGCAG 1129

RESULT 13
US-09-270-767-14085/c
; Sequence 14085, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14085
; LENGTH: 2180
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14085

Query Match          79.0%; Score 15.8; DB 4; Length 2180;
Best Local Similarity 89.5%; Pred. No. 3.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TCCAGCGCGCTCCAGCTG 20
Db      1039 TCGTAGCGCGCTCCAGCTG 1021

RESULT 14
US-10-303-812-1/c
; Sequence 1, Application US/10303812
; Patent No. 6770476
; GENERAL INFORMATION:
; APPLICANT: The Japan Atomic Energy Research Institute
```

```
; TITLE OF INVENTION: Radiation Resistant Bacterium/E. Coli Shuttle Vector
; FILE REFERENCE: 001458.00029
; CURRENT APPLICATION NUMBER: US/10/303,812
; CURRENT FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: JP 2002-46377
; PRIOR FILING DATE: 2002-02-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2467
; TYPE: DNA
; ORGANISM: Deinococcus radiopugnans
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(2467)
; OTHER INFORMATION: Nucleotide sequence of cryptic plasmid pUE30 from Deinococcus ra
US-10-303-812-1

Query Match          79.0%; Score 15.8; DB 4; Length 2467;
Best Local Similarity 89.5%; Pred. No. 3.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ATCCAGCGCGCTCCAGCT 19
Db      1803 ACCCAGCGCTGCTCCAGCT 1785

RESULT 15
US-09-902-540-1084
; Sequence 1084, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1084
; LENGTH: 9035
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1084

Query Match          79.0%; Score 15.8; DB 4; Length 9035;
Best Local Similarity 89.5%; Pred. No. 4.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 TCCAGCGCGCTCCAGCTG 20
Db      1468 TCCCGCGCGCTCCAGCAG 1486

RESULT 16
US-09-949-016-13013
; Sequence 13013, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```


RESULT 21
US-09-252-991A-12304
; Sequence 12304, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12304
; LENGTH: 1632
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12304

Query Match 77.0%; Score 15.4; DB 4; Length 1632;
Best Local Similarity 94.1%; Pred. No. 5.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 CCAGCGCGCTCCAGCTG 20
|||||
DB 225 CCAGCGCGCTCCAGCTG 241

RESULT 22
US-09-902-540-8807/c
; Sequence 8807, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8807
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-8807

Query Match 77.0%; Score 15.4; DB 4; Length 3177;
Best Local Similarity 94.1%; Pred. No. 6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 TCCAGCGCGCTCCAGC 18
|||||
DB 1798 TCCAGCGCGCTCCAGC 1782

RESULT 23
US-09-902-540-891
; Sequence 891, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 891
; LENGTH: 6380
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-891
Query Match 77.0%; Score 15.4; DB 4; Length 6380;
Best Local Similarity 94.1%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 CCAGCGCGCTCCAGCT 19
|||||
DB 1489 CTCAGCGCGCTCCAGCT 1505
RESULT 24
US-08-056-200-93
; Sequence 93, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1507..1644
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1645..2511
; FEATURE:

NAME/KEY: CDS
LOCATION: 2512..8070
US-08-056-200-93

Query Match
Best Local Similarity 77.0%; Score 15.4; DB 1; Length 9551;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGC 18
DB 8006 TCCAGTGGCTCCAGC 8022

RESULT 25
US-08-800-644-93
Sequence 93, Application US/08800644
Patent No. 5958752
GENERAL INFORMATION:
APPLICANT: Steinert, Peter M.
APPLICANT: Lee, Seung-Chul
APPLICANT: Kim, In-Gyu
APPLICANT: Chung, Soo-Il
APPLICANT: Park, Sang-Chul
TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
TITLE OF INVENTION: Methods of Using Same
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive, Sixteenth Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 14-FEB-1997
APPLICATION NUMBER: US/08/800,644
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/056,200
FILING DATE: 30-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fredrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 9551 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1507..1644
FEATURE:
NAME/KEY: intron
LOCATION: 1645..2511
FEATURE:
NAME/KEY: CDS
LOCATION: 2512..8070
US-08-800-644-93

Query Match
77.0%; Score 15.4; DB 2; Length 9551;

Best Local Similarity 94.1%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGC 18
DB 8006 TCCAGTGGCTCCAGC 8022

RESULT 26
US-09-902-540-936/c
Sequence 936, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkie, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 936
LENGTH: 9880
TYPE: DNA
ORGANISM: Myxococcus xanthus
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(9880)
OTHER INFORMATION: unsure at all n locations
US-09-902-540-936

Query Match
77.0%; Score 15.4; DB 4; Length 9880;
Best Local Similarity 94.1%; Pred. No. 6.3e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCCAGCGCGCTCCAGC 18
DB 1800 TCCAGCGTGTCCAGC 1784

RESULT 27
US-09-105-537-30/c
Sequence 30, Application US/09105537A
Patent No. 6265202
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438US1
CURRENT APPLICATION NUMBER: US/09/105,537A
CURRENT FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30
LENGTH: 13842
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-105-537-30

Query Match
77.0%; Score 15.4; DB 3; Length 13842;
Best Local Similarity 94.1%; Pred. No. 6.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAG 17
DB 578 ATCCGCGCGCTCCAG 562

RESULT 28

US-09-105-537-5/c
; Sequence 5, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 77.0%; Score 15.4; DB 3; Length 36778;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAG 17
||||| |||||||
DB 2319 ATCCCGCGCGCTCCAG 2303

RESULT 29

US-09-320-878-19/c
; Sequence 19, Application US/09320878A
; Patent No. 6117659
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/320,878A
; CURRENT FILING DATE: 1999-05-27
; EARLIER FILING DATE: 1998-05-27
; EARLIER FILING DATE: 1998-08-28
; EARLIER FILING DATE: 1998-05-06
; EARLIER FILING DATE: 1997-04-30
; EARLIER FILING DATE: 1997-04-30
; EARLIER FILING DATE: 1999-02-08
; EARLIER FILING DATE: 1998-09-22
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-320-878-19

Query Match 77.0%; Score 15.4; DB 3; Length 38506;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAG 17
||||| |||||||
DB 461 ATCCCGCGCGCTCCAG 445

RESULT 30

US-09-141-908-1/c
; Sequence 1, Application US/09141908
; Patent No. 6503741
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a
; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold
; FILE REFERENCE: 300622002100
; CURRENT APPLICATION NUMBER: US/09/141,908
; CURRENT FILING DATE: 1998-08-28
; EARLIER FILING DATE: 1998-05-06
; EARLIER FILING DATE: 1998-05-06
; EARLIER FILING DATE: 1997-04-30
; EARLIER FILING DATE: 1997-04-30
; EARLIER FILING DATE: 1998-03-05
; EARLIER FILING DATE: 1998-03-05
; EARLIER FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-141-908-1

Query Match 77.0%; Score 15.4; DB 4; Length 38506;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAG 17
||||| |||||||
DB 461 ATCCCGCGCGCTCCAG 445

RESULT 31

US-09-657-440-19/c
; Sequence 19, Application US/09657440
; Patent No. 6509455
; GENERAL INFORMATION:
; APPLICANT: ASHLEY, Gary
; APPLICANT: BETLACH, Melanie C.
; APPLICANT: BETLACH, Mary C.
; APPLICANT: MCDANIEL, Robert
; APPLICANT: TANG, Li
; TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 300622002120
; CURRENT APPLICATION NUMBER: US/09/657,440
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/320,878
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: CIP OF 09/141,908
; PRIOR FILING DATE: 1998-08-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 38506
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-657-440-19

Query Match 77.0%; Score 15.4; DB 4; Length 38506;
Best Local Similarity 94.1%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCCCGCGCGCTCCAG 17
||||| |||||||
DB 461 ATCCCGCGCGCTCCAG 445

RESULT 32
US-09-949-016-13047/c
; Sequence 13047, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13047
; LENGTH: 63183
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(63183)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13047

Query Match 77.0%; Score 15.4; DB 4; Length 63183;
Best Local Similarity 94.1%; Pred. No. 6.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATCCGAGCGGCTCCAG 17
|||||

Db 23194 ATCCGAGCGGCTCCAG 23178

RESULT 33
US-09-949-016-13048/c
; Sequence 13048, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13048
; LENGTH: 63183
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(63183)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13048

Query Match 77.0%; Score 15.4; DB 4; Length 63183;
Best Local Similarity 94.1%; Pred. No. 6.7e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATCCGAGCGGCTCCAG 17
|||||

Db 23194 ATCCGAGCGGCTCCAG 23178

RESULT 34
US-09-513-999C-4006/c
; Sequence 4006, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 4006
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 74..379
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 252..g
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 60
; OTHER INFORMATION: Xaa=His or Arg
US-09-513-999C-4006

Query Match 76.0%; Score 15.2; DB 4; Length 463;
Best Local Similarity 85.0%; Pred. No. 6.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ATCCGAGCGGCTCCAGCTG 20
|||||

Db 336 ATCCATGGAGCTCCAGCTG 317

RESULT 35
US-09-351-150A-40
; Sequence 40, Application US/09351150A
; Patent No. 6538177
; GENERAL INFORMATION:
; APPLICANT: Duwick, J
; TITLE OF INVENTION: Compositions and Methods for Fumonisin Detoxification
; FILE REFERENCE: 035718/167676, 5718-62
; CURRENT APPLICATION NUMBER: US/09/351,150A
; CURRENT FILING DATE: 1998-07-12
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Bacterium 2412.1
; FEATURE:
; OTHER INFORMATION: fcct
; NAME/KEY: CDS
; LOCATION: (1)..(531)
US-09-351-150A-40

Query Match 76.0%; Score 15.2; DB 4; Length 531;
Best Local Similarity 85.0%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
||||| ||||| |||||
Db 320 ATCCCTTCGGCTCGAGCTG 339

RESULT 36

US-09-949-016-39156
; Sequence 39156, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39156
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-39156

Query Match 76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
||||| ||||| |||||
Db 470 ATCCAGGCCCTCCAGCTG 489

RESULT 37

US-09-949-016-112208
; Sequence 112208, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112208
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-112208

Query Match 76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
||||| ||||| |||||
Db 470 ATCCAGGCCCTCCAGCTG 489

RESULT 38

US-09-949-016-112249
; Sequence 112249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112249
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-112249

Query Match 76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
||||| ||||| |||||
Db 470 ATCCAGGCCCTCCAGCTG 489

RESULT 39

US-09-949-016-187754
; Sequence 187754, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187754
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-187754

Query Match 76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATCCAGCGCGCTCCAGCTG 20
||||| ||||| |||||
Db 470 ATCCAGGCCCTCCAGCTG 489

RESULT 40

US-09-949-016-187795
; Sequence 187795, Application US/09949016
; Patent No. 6812339

```

; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 187795
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-187795

```

```

Query Match      76.0%; Score 15.2; DB 4; Length 601;
Best Local Similarity 85.0%; Pred. No. 7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 ATCCGAGCGGCTCCAGCTG 20
         |||||
DB      470 ATCCGAGCGGCTCCAGCTG 489

```

```

Search completed: June 4, 2005, 11:53:09
Job time : 54.497 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 04:22:34 ; Search time 166.757 Seconds
(without alignments)
780.981 Million cell updates/sec

Title: US-09-674-277-24

Perfect score: 22

Sequence: 1 acccatgatggcgcattgatg 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	3	Aaz36124 Primer de
2	22	100.0	1181	3	Aaz36102 Nucleic a
3	17.8	80.9	723	9	Adb07361 Allostoc
4	17.8	80.9	723	9	Adb07359 Allostoc
5	17.8	80.9	9570	9	Adb07375 Allostoc
6	17.8	80.9	9570	9	Adb07377 Allostoc
7	17.8	80.9	29584	12	Adg97635 Human can
8	17.8	80.9	110000	9	Continuation (3 of
9	17.8	80.9	110000	9	Continuation (4 of
10	17.4	79.1	1569	4	Abi26449 Drosophil
11	17.4	79.1	4937	4	Abi26448 Drosophil
12	17.2	78.2	672	11	Ach95112 Klebsiell
13	17.2	78.2	975	12	Adi30088 M. methyl
14	17.2	78.2	1823	13	Adr21932 Methyloph
15	17.2	78.2	37051	8	Aad56126 Mouse Noc
16	17.2	78.2	37051	9	Ada02489 Mouse Noc
17	17.2	78.2	37051	10	Adb72227 Mouse Noc
18	17.2	78.2	100610	12	Adg97608 Mouse can
19	16.8	76.4	312	11	Ach96947 Klebsiell
20	16.8	76.4	359	4	Aak88001 Human dig

c	21	16.8	76.4	659	12	ADJ40155
	22	16.8	76.4	759	11	ACH97207
	23	16.8	76.4	1359	3	AAC50463
	24	16.8	76.4	1363	3	AAC36643
	25	16.8	76.4	1504	4	ABL05649
	26	16.8	76.4	1629	13	ADT43185
	27	16.8	76.4	8018	4	ABL05648
	28	16.8	76.4	11197	6	ABK12708
	29	16.8	76.4	73882	13	ADG73531
c	30	16.4	74.5	316	6	ABL67308
	31	16.4	74.5	316	6	ABN94178
	32	16.4	74.5	539	4	AAH11562
	33	16.4	74.5	600	6	ABK34361
	34	16.4	74.5	852	6	ABK73908
	35	16.4	74.5	1266	3	AAF18110
c	36	16.4	74.5	1712	4	AAH98188
	37	16.4	74.5	1773	3	AAZ43799
	38	16.4	74.5	1984	6	ABL55085
	39	16.4	74.5	1984	8	ADA39911
	40	16.4	74.5	1984	10	ADA56100
	41	16.4	74.5	2065	4	AAS34816
	42	16.4	74.5	2065	10	ADC45974
	43	16.4	74.5	2229	3	AAZ22234
	44	16.4	74.5	2552	11	ADM01705
	45	16.4	74.5	2614	10	ADC99151

ALIGNMENTS

RESULT 1

AZ36124

ID AZ36124 standard; DNA; 22 BP.

XX

AC AZ36124;

DT 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

DE Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

XX WO9955908-A2.

PN 04-NOV-1999.

PD 27-APR-1999; 99WO-FR001000.

XX 28-APR-1999; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PI Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic

PT Escherichia coli, particularly serotype O157:H7, used for detecting these

PT bacteria in food.

XX Claim 5; Page 27; 48pp; French.

XX AZ36103-27 represent fragments derived from nucleic acid sequences specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are derived from two sequences. The first (AAZ36101) is 99.9% homologous to the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102). The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AA236102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of *Shigella flexneri*. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX SQ Sequence 22 BP; 5 A; 6 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 3; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.38; Mismatches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0

QY 1 ACCCATGATGGCGCATCTGATG 22
 |||||
 DB 1 ACCCATGATGGCGCATCTGATG 22

RESULT 2

AA236102/c
 ID AA236102 standard; DNA; 1181 BP.

XX AC AA236102;

XX DT 11-FEB-2000 (first entry)

XX DE Nucleic acid sequence specific to enterohemorrhagic *Escherichia coli*.

XX KW Enterohemorrhagic *Escherichia coli*; EHEC; virulence factor;

XX KW enterohemolysine; ehly; intimin; eae; virK gene; *E. coli* O157:H7; db.

XX OS *Escherichia coli*.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOPI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX PS WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
 XX *Escherichia coli*, particularly serotype O157:H7, used for detecting these
 XX bacteria in food.

XX Claim 1; Fig 2; 48pp; French.

XX CC The present sequence is specific to enterohemorrhagic *Escherichia coli*
 CC (EHEC). The sequence associated with the presence of virulence factors
 CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
 CC 68% homology with the virK gene which codes for virulence proteins of
 CC *Shigella flexneri*. The present sequence is of plasmid origin. Fragments
 CC of the present sequence are used, as probes and primers, for detection of
 CC *E. coli* O157:H7 and other enterohemorrhagic *E. coli* (EHEC), in human or
 CC animal samples, foods or the environment. The fragments are also useful
 CC for epidemiological studies

XX SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 3; Length 1181;
 Best Local Similarity 100.0%; Pred. No. 0.79;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
 |||||
 DB 884 ACCCATGATGGCGCATCTGATG 863

RESULT 3

ADB07361/c

ID ADB07361 standard; DNA; 723 BP.

XX AC ADB07361;

XX DT 20-NOV-2003 (first entry)

XX DE *Alloiooccus otitis* antigenic protein encoding DNA SEQ ID NO:1301.

XX KW *Alloiooccus otitis*; antigenic protein; immunogenic; immunisation;
 KW gene therapy; Gram-positive bacterium; infection; gene; ds.

XX OS *Alloiooccus otitis*.

XX PN WO2003048304-A2.

XX PD 12-JUN-2003.

XX PF 25-NOV-2002; 2002WO-US036123.

XX PR 29-NOV-2001; 2001US-0333777P.

XX PR 18-NOV-2002; 2002US-0426742P.

XX PA (AMHP) WYETH HOLDINGS CORP.

XX PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;

XX WPI; 2003-505284/47.

XX P-PSDB; ADB07362.

XX PT New *Alloiooccus otitis* polynucleotides and polypeptides, useful for
 XX treating and diagnosing diseases, drug screening assays and monitoring of
 XX effects during drug clinical trials.

XX PS Claim 7; SEQ ID NO 1301; 1019pp; English.

XX CC The present invention describes an isolated polynucleotide (1) of
 CC *Alloiooccus otitis* genomic DNA, which encodes an antigenic protein.
 CC *Alloiooccus otitis* is a Gram-positive bacterium. Also described: (1)
 CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an
 CC expression vector comprising the novel isolated polynucleotide (1), its
 CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC host cell, transformed, transformed or infected with the vector of (2);
 CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC composition comprising the polypeptide, its complement, biological
 CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC expression vector; (6) a pharmaceutical composition comprising the
 CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC immunising against *Alloiooccus otitis* by administering to a host the
 CC immunogenic composition; (9) detecting and/or identifying *Alloiooccus*
 CC *otitis* in the biological sample; (10) a kit comprising a container
 CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC or the antibody of (4); and (11) producing a polypeptide by culturing the
 CC genetically engineered host cell under conditions suitable to produce the
 CC polypeptide from the culture. (1) can be used in gene therapy. The
 CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC invention can be used for treating and diagnosing diseases, drug
 CC screening assays and monitoring of effects during drug clinical trials.
 CC The polynucleotides are useful for expressing and detecting *Alloiooccus*
 CC *otitis*. The present sequence encodes an *Alloiooccus otitis* antigen
 CC protein from the present invention.

XX SQ Sequence 723 BP; 154 A; 194 C; 159 G; 216 T; 0 U; 0 Other;

Query Match 80.9%; Score 17.8; DB 9; Length 723;
 Best Local Similarity 90.5%; Pred. No. 92;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCATCTGAT 21

Db 183 ACCACGATGGCGCAACTGAT 163
||||| ||||||| |||||||

RESULT 4
ADB07359/c
ID ADB07359 standard; DNA; 723 BP.
XX AC ADB07359;
XX DT 20-NOV-2003 (first entry)
XX DE Alloiococcus otitis antigenic protein encoding DNA SEQ ID NO:1299.
XX KW Alloiococcus otitidis; antigenic protein; immunogenic; immunisation;
XX KW gene therapy; Gram-positive bacterium; infection; gene; ds.
XX OS Alloiococcus otitis.
XX PN WO2003048304-A2.
XX PD 12-JUN-2003.
XX PF 25-NOV-2002; 2002WO-US036123.
XX PR 29-NOV-2001; 2001US-0333777P.
XX PR 18-NOV-2002; 2002US-0426742P.
XX PA (AMHP) WYETH HOLDINGS CORP.
XX PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX DR WPI; 2003-505284/47.
XX DR P-PSDB; ADB07360.
XX PT New Alloiococcus otitidis polynucleotides and polypeptides, useful for
XX PT treating and diagnosing diseases, drug screening assays and monitoring of
XX PT effects during drug clinical trials.
XX PS Claim 7; SEQ ID NO 1299; 1019pp; English.
XX CC The present invention describes an isolated polynucleotide (1) of
XX CC Alloiococcus otitidis genomic DNA, which encodes an antigenic protein.
XX CC Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1)
XX CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an
XX CC expression vector comprising the novel isolated polynucleotide (1), its
XX CC complement, degenerate variant or fragment; (3) a genetically engineered
XX CC host cell, transfected, transformed or infected with the vector of (2);
XX CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
XX CC composition comprising the polypeptide, its complement, biological
XX CC equivalent or fragment, or the polynucleotide that is comprised in the
XX CC expression vector; (6) a pharmaceutical composition comprising the
XX CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
XX CC of the polypeptides of (1), their biological equivalent or fragment; (8)
XX CC immunising against Alloiococcus otitidis by administering to a host the
XX CC otitidis in the biological sample; (10) a kit comprising Alloiococcus
XX CC containing the novel polynucleotide, its degenerate variant or fragment,
XX CC or the antibody of (4); and (11) producing a polypeptide by culturing the
XX CC polypeptide from the culture. (1) can be used in gene therapy. The
XX CC polynucleotides, polypeptides, antibodies and compositions of the present
XX CC invention can be used for treating and diagnosing diseases, drug
XX CC screening assays and monitoring of effects during drug clinical trials.
XX CC The polynucleotides are useful for expressing and detecting Alloiococcus
XX CC otitidis. The present sequence encodes an Alloiococcus otitidis antigen
XX CC protein from the present invention.
XX SQ Sequence 723 BP; 154 A; 194 C; 159 G; 216 T; 0 U; 0 Other;

Query Match 80.9%; Score 17.8; DB 9; Length 723;
Best Local Similarity 90.5%; Pred. No. 92;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||||| |||||||
Db 183 ACCACGATGGCGCAACTGAT 163

RESULT 5
ADB07375
ID ADB07375 standard; DNA; 9570 BP.
XX AC ADB07375;
XX DT 20-NOV-2003 (first entry)
XX DE Alloiococcus otitis antigenic protein encoding DNA SEQ ID NO:1315.
XX KW Alloiococcus otitidis; antigenic protein; immunogenic; immunisation;
XX KW gene therapy; Gram-positive bacterium; infection; gene; ds.
XX OS Alloiococcus otitis.
XX PN WO2003048304-A2.
XX PD 12-JUN-2003.
XX PF 25-NOV-2002; 2002WO-US036123.
XX PR 29-NOV-2001; 2001US-0333777P.
XX PR 18-NOV-2002; 2002US-0426742P.
XX PA (AMHP) WYETH HOLDINGS CORP.
XX PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
XX DR WPI; 2003-505284/47.
XX DR P-PSDB; ADB07376.
XX PT New Alloiococcus otitidis polynucleotides and polypeptides, useful for
XX PT treating and diagnosing diseases, drug screening assays and monitoring of
XX PT effects during drug clinical trials.
XX PS Claim 7; SEQ ID NO 1315; 1019pp; English.
XX CC The present invention describes an isolated polynucleotide (1) of
XX CC Alloiococcus otitidis genomic DNA, which encodes an antigenic protein.
XX CC Alloiococcus otitidis is a Gram-positive bacterium. Also described: (1)
XX CC an isolated polypeptide that is encoded by the polynucleotide (1); (2) an
XX CC expression vector comprising the novel isolated polynucleotide (1), its
XX CC complement, degenerate variant or fragment; (3) a genetically engineered
XX CC host cell, transfected, transformed or infected with the vector of (2);
XX CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
XX CC composition comprising the polypeptide, its complement, biological
XX CC equivalent or fragment, or the polynucleotide that is comprised in the
XX CC expression vector; (6) a pharmaceutical composition comprising the
XX CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
XX CC of the polypeptides of (1), their biological equivalent or fragment; (8)
XX CC immunising against Alloiococcus otitidis by administering to a host the
XX CC otitidis in the biological sample; (10) a kit comprising Alloiococcus
XX CC containing the novel polynucleotide, its degenerate variant or fragment,
XX CC or the antibody of (4); and (11) producing a polypeptide by culturing the
XX CC polypeptide from the culture. (1) can be used in gene therapy. The
XX CC polynucleotides, polypeptides, antibodies and compositions of the present
XX CC invention can be used for treating and diagnosing diseases, drug
XX CC screening assays and monitoring of effects during drug clinical trials.
XX CC The polynucleotides are useful for expressing and detecting Alloiococcus
XX CC otitidis. The present sequence encodes an Alloiococcus otitidis antigen
XX CC protein from the present invention.
XX SQ Sequence 9570 BP; 3048 A; 2117 C; 2312 G; 2093 T; 0 U; 0 Other;

Query Match 80.9%; Score 17.8; DB 9; Length 9570;

Best Local Similarity 90.5%; Pred. No. 1.5e+02; Mismatches 2; Indels 0; Gaps 0;
Matches 19; Conservative 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||||| |||||
Db 5668 ACCCAGATGGCGCAACTGAT 5688

RESULT 6
ADB07377
ID ADB07377 standard; DNA; 9570 BP.
XX
AC ADB07377;
XX
DT 20-NOV-2003 (first entry)
XX
DE Allolococcus otitis antigenic protein encoding DNA SEQ ID NO:1317.
XX
DE Allolococcus otitis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection; gene; ds.
XX
OS Allolococcus otitis.
XX
PN WO2003048304-A2.
XX
PD 12-JUN-2003.
XX
PF 25-NOV-2002; 2002WO-US036123.
XX
PR 29-NOV-2001; 2001US-0333777P.
PR 18-NOV-2002; 2002US-0426742P.
XX
PA (AMHP) WYETH HOLDINGS CORP.
XX

PI Fletcher LD, Mcmichael JC, Russell DP, Zagursky RJ;
XX
DR WPI: 2003-505284/47.
XX
DR P-PSDB; ADB07378.
XX
PT New Allolococcus otitis polynucleotides and polypeptides, useful for
PT treating and diagnosing diseases, drug screening assays and monitoring of
PT effects during drug clinical trials.
XX
PS Claim 7; SEQ ID NO 1317; 1019pp; English.

XX The present invention describes an isolated polynucleotide (I) of
CC Allolococcus otitis genomic DNA, which encodes an antigenic protein.
CC Allolococcus otitis is a Gram-positive bacterium. Also described: (1)
CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
CC expression vector comprising the novel isolated polynucleotide (I), its
CC complement, degenerate variant or fragment; (3) a genetically engineered
CC host cell, transfected, transformed or infected with the vector of (2);
CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
CC composition comprising the polypeptide, its complement, biological
CC equivalent or fragment, or the polynucleotide that is comprised in the
CC expression vector; (6) a pharmaceutical composition comprising the
CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
CC of the polypeptides of (1), their biological equivalent or fragment; (8)
CC immunising against Allolococcus otitis by administering to a host the
CC immunogenic composition; (9) detecting and/or identifying Allolococcus
CC otitis in the biological sample; (10) a kit comprising a container
CC containing the novel polynucleotide, its degenerate variant or fragment,
CC or the antibody of (4); and (11) producing a polypeptide by culturing the
CC genetically engineered host cell under conditions suitable to produce the
CC polypeptide from the culture. (I) can be used in gene therapy. The
CC polynucleotides, polypeptides, antibodies and compositions of the present
CC invention can be used for treating and diagnosing diseases, drug
CC screening assays and monitoring of effects during drug clinical trials.
CC The polynucleotides are useful for expressing and detecting Allolococcus
CC otitis. The present sequence encodes an Allolococcus otitis antigen
XX protein from the present invention.

XX Sequence 9570 BP; 3048 A; 2117 C; 2312 G; 2093 T; 0 U; 0 Other;

Query Match 80.9%; Score 17.8; DB 9; Length 9570;
Best Local Similarity 90.5%; Pred. No. 1.5e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||||| |||||
Db 5668 ACCCAGATGGCGCAACTGAT 5688

RESULT 7
ADQ97635/c
ID ADQ97635 standard; DNA; 29584 BP.
XX
AC ADQ97635;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human cancer associated sequence HD10-019, SEQ ID 612.
XX
KW Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human; ds.
XX
OS Homo sapiens.
XX
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI: 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 612; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 29584 BP; 7268 A; 6620 C; 6719 G; 7207 T; 0 U; 1770 Other;

Query Match 80.9%; Score 17.8; DB 12; Length 29584;
Best Local Similarity 90.5%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||||| |||||
Db 15324 ACCCATGATGGCGTTCTGAT 15304

RESULT 8
ADB12064_02
Continuation (3 of 18) of ADB12064 from base 200001 (Allolococcus otitis entire genome
WP Sequence split into 18 fragments LOCUS ADB12064 Accession Adb12064
WP Fragment Name Begin End
WP ADB12064_00 1 110000
WP ADB12064_01 100001 210000
WP ADB12064_02 200001 310000
WP ADB12064_03 300001 410000
WP ADB12064_04 400001 510000
WP ADB12064_05 500001 610000

WP ADB12064_06 600001 710000
WP ADB12064_07 700001 810000
WP ADB12064_08 800001 910000
WP ADB12064_09 900001 1010000
WP ADB12064_10 1000001 1110000
WP ADB12064_11 1100001 1210000
WP ADB12064_12 1200001 1310000
WP ADB12064_13 1300001 1410000
WP ADB12064_14 1400001 1510000
WP ADB12064_15 1500001 1610000
WP ADB12064_16 1600001 1710000
WP ADB12064_17 1700001 1754382

Query Match 80.9%; Score 17.8; DB 9; Length 110000;
Best Local Similarity 90.5%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCACTGTAT 21
||||| ||||||| |||||||

Db 105407 ACCCAGTGGCGCACTGTAT 105427

RESULT 9

ADBI2064_03
Continuation (4 of 18) of ADBI2064 from base 300001 (Alloiococcus otitis entire genome
WP Sequence split into 18 fragments LOCUS ADBI2064 Accession ADBI2064

Fragment Name	Begin	End
WP ADBI2064_00	1	110000
WP ADBI2064_01	100001	210000
WP ADBI2064_02	200001	310000
WP ADBI2064_03	300001	410000
WP ADBI2064_04	400001	510000
WP ADBI2064_05	500001	610000
WP ADBI2064_06	600001	710000
WP ADBI2064_07	700001	810000
WP ADBI2064_08	800001	910000
WP ADBI2064_09	900001	1010000
WP ADBI2064_10	1000001	1110000
WP ADBI2064_11	1100001	1210000
WP ADBI2064_12	1200001	1310000
WP ADBI2064_13	1300001	1410000
WP ADBI2064_14	1400001	1510000
WP ADBI2064_15	1500001	1610000
WP ADBI2064_16	1600001	1710000
WP ADBI2064_17	1700001	1754382

Query Match 80.9%; Score 17.8; DB 9; Length 110000;
Best Local Similarity 90.5%; Pred. No. 2.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCACTGTAT 21
||||| ||||||| |||||||

Db 5407 ACCCAGTGGCGCACTGTAT 5427

RESULT 10

ADBI206449
ID ABL26449 standard; DNA; 1569 BP.

AC ABL26449;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 30820.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEXE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 30820; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABL30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABBS57737-
CC ABBY2072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1569 BP; 304 A; 520 C; 454 G; 291 T; 0 U; 0 Other;

Query Match 79.1%; Score 17.4; DB 4; Length 1569;
Best Local Similarity 94.7%; Pred. No. 1.7e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTG 19
||||| ||||||| |||||||
Db 264 ACCCATGATGGCGCATCTG 282

RESULT 11
ADBI26448/c
ID ABL26448 standard; DNA; 4937 BP.
AC ABL26448;
XX
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 30817.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEXE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 30817; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIFO at ftp.wifo.int/pub/published_pct_sequences
XX
SQ Sequence 4937 BP; 1362 A; 1099 C; 1171 G; 1305 T; 0 U; 0 Other;

Query Match 79.1%; Score 17.4; DB 4; Length 4937;
Best Local Similarity 94.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTG 19
Db 3402 ACCCATGATGGCGCATCTG 3384

RESULT 12
ACH95112
ID ACH95112 standard; DNA; 672 BP.
AC ACH95112;
XX
XX 29-JUL-2004 (first entry)
DE Klebsiella pneumoniae polynucleotide seqid 907.
XX
XX Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; vaccine; gene; ds.
XX
XX Klebsiella pneumoniae.
XX
XX US6610836-B1.
XX
XX 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
XX
XX 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI: 2003-895346/82.
XX P-PSDB; ABO61561.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 907; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This sequence encodes a Klebsiella
XX pneumoniae polypeptide of the invention
XX
SQ Sequence 672 BP; 136 A; 169 C; 224 G; 143 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 11; Length 672;
Best Local Similarity 86.4%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
Db 300 ACCGATGCTGGCGCAGCTGATG 321

RESULT 13
ADI30088
ID ADI30088 standard; DNA; 975 BP.
XX
XX ADI30088;
XX
XX 22-APR-2004 (first entry)
XX
XX M. methylotrophus methenyl H4 MPT cyclohydrolase gene.
XX
XX Cl assimilation; L-amino acid production; fermentation; gene; ds.
XX
XX Methylophilus methylotrophus.
XX
XX Key Location/Qualifiers
XX CDS 1..975
XX FT /*tag= a
XX FT /product= "M. methylotrophus enzyme involved in Cl
XX FT assimilation"
XX
XX US2003232338-A1.
XX
XX 18-DEC-2003.
XX
XX 12-JUN-2002; 2002US-00166653.
XX
XX 12-JUN-2002; 2002US-00166653.
XX
XX (USUD/) USUDA Y.
XX (NISH/) NISHIO Y.
XX (YASU/) YASUEDA H.
XX (SUGI/) SUGIMOTO S.
XX
XX Usuda Y, Nishio Y, Yasueda H, Sugimoto S;
XX
XX WPI: 2004-061268/06.
XX P-PSDB; ADI30089.
XX
XX New polynucleotides encoding polypeptides involved in one-carbon compound
XX metabolism in *Methylophilus methylotrophus*, useful for producing L-amino
XX acids (e.g. threonine or serine) or as probes and/or as primers.
XX
XX Claim 13; SEQ ID NO 5; 53pp; English.
XX
XX The invention relates to enzymes involved in Cl assimilation in
XX *Methylophilus methylotrophus* and their corresponding polynucleotides. The
XX invention also relates to methods of producing amino acids in micro-
XX organisms having enhanced or attenuated expression of the polypeptides
XX and polynucleotides involved in Cl assimilation. The invention is useful
XX for producing L-amino acids (e.g. asparagine, threonine or serine) using
XX *Methylophilus bacteria*. It is also used in identifying microorganisms
XX that can be used to produce fine chemicals, for e.g. by fermentative
XX processes. The present sequence is M. methylotrophus gene encoding an
XX enzyme involved in Cl assimilation.
XX
SQ Sequence 975 BP; 228 A; 255 C; 250 G; 242 T; 0 U; 0 Other;

Query Match 78.2%; Score 17.2; DB 12; Length 975;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
Db 54 ACCACTGCTGGCGCATCTGATG 75

RESULT 14

```
ADR21932
ID ADR21932 standard; DNA; 1823 BP.
XX
AC ADR21932;
XX
DT 04-NOV-2004 (first entry)
XX
DE Methylophilus methylophilus mch coding sequence.
XX
KW recombinant methanol-assimilating microorganism; gene; ds; mch.
XX
OS Methylophilus methylophilus.
XX
FH Key Location/Qualifiers
FT CDS 522..1496
FT /tag= a
FT /product= "Methylophilus methylophilus mch protein"
XX
XX JP2004229662-A.
XX
XX 19-AUG-2004.
XX
XX 07-JAN-2004; 2004JP-00001717.
XX
XX 08-JAN-2003; 2003JP-00001927.
XX
XX (AJIN ) AJINOMOTO KK.
XX
XX WPI; 2004-586588/57.
XX
XX P-PSDB; ADR21933.
XX
XX Generating recombinant methanol-assimilating microorganism, by
PT introducing DNA fragment with base sequence similar to base sequence of
PT arbitrary regions in chromosomal DNA, into methanol-assimilating
PT microorganism.
XX
XX Example 3; SEQ ID NO 23; 10pp; Japanese.
XX
XX The invention comprises a method for generating recombinant methanol-
CC assimilating microorganisms. The method involves introducing foreign DNA
CC into the chromosomal DNA of a methanol-assimilating microorganism, by
CC introducing a linear DNA fragment having a base sequence similar to the
CC base sequence of arbitrary regions in the chromosomal DNA of the methanol
CC -assimilating microorganism, and selecting the recombinant microorganism.
CC The method of the invention is useful for generating methanol-
CC assimilating microorganisms. The present DNA sequence encodes the
CC Methylophilus methylophilus mch protein.
XX
XX Sequence 1823 BP; 437 A; 459 C; 484 G; 443 T; 0 U; 0 Other;
SQ
Query Match 78.2%; Score 17.2; DB 13; Length 1823;
Best Local Similarity 86.4%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 ACCCATGATGGGCATCTGATG 22
DB 575 ACCACTGGTGGGCATCTGATG 596
RESULT 15
AAD56126
ID AAD56126 standard; DNA; 37051 BP.
XX
AC AAD56126;
XX
DT 07-AUG-2003 (first entry)
XX
DE Mouse Notch carcinoma associated (CA) gene.
XX
KW Carcinoma; gene therapy; carcinoma associated gene; CA; Notch; mouse; ds.
XX
XX Mus sp.
OS
XX
PN WO2003035837-A2.
XX
PD 01-MAY-2003.
XX
PF 22-OCT-2002; 2002WO-US033835.
XX
PR 23-OCT-2001; 2001US-00004113.
XX
PA (SAGR-) SAGRES DISCOVERY.
XX
PI Engelhard EK, Morris DW;
XX
DR WPI; 2003-421412/39.
XX
XX New recombinant nucleic acid and its encoded protein, useful for
PT preparing a composition for diagnosing or treating carcinomas.
XX
XX Claim 1; Page 157-162; 173pp; English.
XX
XX The invention relates to novel sequences which are useful for preparing a
CC composition for diagnosing or treating carcinomas. These sequence are
CC also useful in gene therapy. The present sequence is mouse Notch
CC carcinoma associated (CA) gene. This sequence is used in the invention
XX
XX Sequence 37051 BP; 8610 A; 9436 C; 9720 G; 8760 T; 0 U; 525 Other;
SQ
Query Match 78.2%; Score 17.2; DB 8; Length 37051;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 ACCCATGATGGGCATCTGATG 22
DB 27953 ACCCAGGATGGCCCATCTGCTG 27974
RESULT 16
ADA02489
ID ADA02489 standard; DNA; 37051 BP.
XX
AC ADA02489;
XX
DT 06-NOV-2003 (first entry)
XX
DE Mouse Notchl carcinoma associated gene, SEQ ID NO:1007.
XX
KW Mouse; murine; carcinoma associated; oncogene; carcinoma; cancer; breast;
KW prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
KW gene; ds.
XX
XX Mus sp.
OS
XX
XX WO2003057146-A2.
XX
PN 17-JUL-2003.
XX
XX 26-DEC-2002; 2002WO-US041414.
XX
XX 26-DEC-2001; 2001US-00035832.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
PI
XX
XX WPI; 2003-587068/55.
XX
XX New recombinant nucleic acid encoding carcinoma associated protein,
PT useful for preparing compositions for treating carcinomas.
XX
XX Claim 1; SEQ ID NO 1007; 245pp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
CC acid sequences from mouse and human (ADA01482-ADA03094), and to
CC recombinant carcinoma associated proteins (CAP) encoded by them. The
CC
```

CC invention also encompasses expression vectors and host cells comprising a
CC CA nucleic acid, a polypeptide (especially an antibody) that specifically
CC binds to the protein, and a biochip comprising CA nucleic acid or
CC fragments thereof. The sequences of the invention were identified using
CC oncogenic retroviruses, which insert into the genome of the host organism
CC at random. Many of these do not carry transduced host oncogenes or
CC pathogenic trans-acting viral genes, meaning that cancer incidence is a
CC direct consequence of the effects of proviral integration into host
CC protooncogenes. The CA nucleic acid sequences can be used to diagnose
CC carcinoma (especially breast cancer, prostate cancer, lymphoma or
CC leukaemia) or a propensity to carcinoma by determination of the sequence
CC of a CA gene, or by determination of CA gene expression in particular
CC tissues. CA nucleic acids, proteins and antibodies are also useful as
CC therapeutic agents and in screening and evaluating drug candidates. The
CC present sequence represents a specifically claimed murine CA nucleic acid
CC sequence of the invention. Note: The complete sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 37051 BP; 8610 A; 9436 C; 9720 G; 8760 T; 0 U; 525 Other;
Query Match 78.2%; Score 17.2; DB 9; Length 37051;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
||||| ||||| ||||| ||||| |||||
Db 27953 ACCCAGATGGCCCATCTGCTG 27974

RESULT 17
ADB72227
ID ADB72227 standard; DNA; 37051 BP.
XX
AC ADB72227;
XX
DT 04-DEC-2003 (first entry)
XX
DE Mouse Notch1 gene.
XX
KW mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.

OS Mus sp.
XX
XX WO2003008583-A2.
XX
XX 30-JAN-2003.

XX 26-DEC-2001; 2001WO-US051291.
XX
XX 02-MAR-2001; 2001US-00798586.
PR 23-OCT-2001; 2001US-00004113.
PR 08-NOV-2001; 2001US-00052482.
PR 30-NOV-2001; 2001US-00997722.
PR 20-DEC-2001; 2001US-00034650.

XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW, Engelhard EK;
XX
XX WPI; 2003-239337/23.

XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
PT cancers, neoplasm, adenocarcinoma, or sarcomas.

PS Claim 1; SEQ ID NO 55; 2304pp; English.

XX The invention relates to a novel recombinant nucleic acid comprising a
CC nucleotide sequence selected from any of the 660 sequences fully defined
CC in the specification. A polynucleotide of the invention has cytostatic
CC activity, and may have a use in gene therapy, or in a vaccine. The

CC recombinant nucleic acids and polypeptides are useful for treating
CC carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
CC sarcomas. The present sequence represents a mouse gene of the invention.

SQ Sequence 37051 BP; 8610 A; 9436 C; 9720 G; 8760 T; 0 U; 525 Other;
Query Match 78.2%; Score 17.2; DB 10; Length 37051;
Best Local Similarity 86.4%; Pred. No. 3.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
||||| ||||| ||||| ||||| |||||
Db 27953 ACCCAGATGGCCCATCTGCTG 27974

RESULT 18
ADQ97608
ID ADQ97608 standard; DNA; 100610 BP.
XX
XX ADQ97608;
XX
XX 07-OCT-2004 (first entry)
XX
XX Mouse cancer associated sequence MD10-014, SEQ ID 585.

XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse; ds.
XX Mus musculus.
XX WO2004060304-A2.
XX
XX 22-JUL-2004.

XX 22-DEC-2003; 2003WO-US041389.
XX
XX 27-DEC-2002; 2002US-00330773.

XX (SAGR-) SAGRES DISCOVERY INC.
XX
XX Morris DW, Malandro MS;
XX
XX WPI; 2004-543781/52.

XX New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.

PS Claim 1; SEQ ID NO 585; 199pp; English.

XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 100610 BP; 24606 A; 20295 C; 20284 G; 23942 T; 0 U; 11483 Other;
Query Match 78.2%; Score 17.2; DB 12; Length 100610;
Best Local Similarity 86.4%; Pred. No. 4.5e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
||||| ||||| ||||| ||||| |||||
Db 14824 ACTCAAGATGGCGCATCTGATG 14845

RESULT 19
ACH96947/c
ID ACH96947 standard; DNA; 312 BP.
XX
XX ACH96947;
AC
XX

DT 29-JUL-2004 (first entry)
XX Klebsiella pneumoniae polynucleotide seqid 2742.
DE
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
XX
XX
OS Klebsiella pneumoniae.
XX
XX US6610836-B1.
XX
XX 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
XX
XX 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
XX P-PSDB; ABO63396.
XX
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 2742; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This sequence encodes a Klebsiella
XX pneumoniae polypeptide of the invention
XX
XX Sequence 312 BP; 79 A; 116 C; 87 G; 30 T; 0 U; 0 Other;
SQ

Query Match 76.4%; Score 16.8; DB 11; Length 312;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCATGATGGCGATCTGATG 22
|||||
Db 30 CCATGCTGGCGATCTGATG 11

RESULT 20
AAK88001/c
ID AAK88001 standard; cDNA; 359 BP.
XX
XX AAK88001;
XX
XX 05-NOV-2001 (first entry)
DT
XX
XX
DE Human digestive system antigen coding sequence SEQ ID NO: 317.
XX
XX Human; digestive system antigen; Gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200155314-A2.
XX
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001324.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-502630/55.
XX P-PSDB; AAW92228.
XX Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX Claim 1; SEQ ID NO 317; 986pp; English.
PS

XX The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a cDNA encoding a digestive
CC system antigen of the invention
XX
SQ Sequence 359 BP; 85 A; 89 C; 91 G; 90 T; 0 U; 4 Other;
Query Match 76.4%; Score 16.8; DB 4; Length 359;
Best Local Similarity 81.8%; Pred. No. 2.6e+02;
Matches 18; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCCATGATGGCGCATCTGATG 22
Db 125 ACCCAYGATGCGCCTATGATG 104
RESULT 21
ADJ40155/c
ID ADJ40155 standard; cDNA; 659 BP.
XX AC ADJ40155;
XX 06-MAY-2004 (first entry)
XX DE Plant cDNA #1155.
XX Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.
XX
OS Eukaryota.
XX US2004016025-A1.
XX 22-JAN-2004.
XX 26-SEP-2002; 2002US-00260238.
XX 26-SEP-2001; 2001US-0325277P.
XX 26-SEP-2001; 2001US-0325448P.
XX 04-APR-2002; 2002US-0370620P.
XX (BUDW/) BUDWORTH P.
PA (MOUG/) MOUGHAMER T.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZEBROOK J.
PA (GOFF/) GOFF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (PROV/) PROVART N.
PA (RICK/) RICHKE D.
PA (ZHUT/) ZHU T.
XX
PI Budworth P, Moughamer T, Briggs SP, Cooper B, Glazebrook J;
PI Goff SA, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX WPI: 2004-190374/18.
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
XX or high nutritional value.
PS Claim 25; SEQ ID NO 1155; 230pp; English.
XX

CC The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential
CC or constitutive transcription of an operatively linked nucleic acid
CC segment. The invention also relates to a method for augmenting a plant
CC genome and a method of identifying a gene, where its expression is
CC altered in the seed, leaf, stem, panicle, pollen, root or is constitutive
CC in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower,
CC canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley,
CC sorghum, rice or wheat. The polynucleotides and the polypeptides they
CC encode are useful for manipulating crop plants to alter or improve
CC phenotypic characteristics, to produce large quantities of oil or
CC proteins, to incur resistance to insecticides, viruses or fungi, and to
CC incur stress tolerance (e.g. salt, cold or drought) to ensure the plants
CC have a high nutritional value with reduced apical dominance or dwarfism,
CC early flowering or altered metabolic pathways. This sequence represents a
CC plant nucleic acid of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification but was obtained in
CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 659 BP; 141 A; 188 C; 200 G; 129 T; 0 U; 1 Other;

Query Match 76.4%; Score 16.8; DB 12; Length 659;
Best Local Similarity 90.0%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTGAT 21
||||| |||||||
DB 555 CCCATGATGGCGCATCTGAT 536

RESULT 22
ACH97207
ID ACH97207 standard; DNA; 759 BP.
AC ACH97207;
XX
DT 29-JUL-2004 (first entry)
XX
DE Klebsiella pneumoniae polynucleotide seqid 3002.
XX
KW Recombinant expression vector; transcription regulatory element;
KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.
OS Klebsiella pneumoniae.
XX
XX US6610836-B1.
XX
XX 26-AUG-2003.
XX
XX 27-JAN-2000; 2000US-00489039.
XX
XX 29-JAN-1999; 99US-0117747P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Breton GL, Osborne M;
XX
XX WPI; 2003-895346/82.
XX
XX P-PSDB; ABO63656.
XX
XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX
XX Disclosure; SEQ ID NO 3002; 932pp; English.
XX
XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This sequence encodes a Klebsiella
XX pneumoniae polypeptide of the invention

SQ Sequence 759 BP; 107 A; 209 C; 266 G; 177 T; 0 U; 0 Other;
Query Match 76.4%; Score 16.8; DB 11; Length 759;
Best Local Similarity 90.0%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CCCATGATGGCGCATCTGATG 22
||||| |||||||
DB 545 CCCATGATGGCGCATCTGATG 564
RESULT 23
AAC50463
ID AAC50463 standard; DNA; 1359 BP.
XX
AC AAC50463;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 64937.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX
XX 05-MAR-1999; 99US-0123180P.
XX
XX 09-MAR-1999; 99US-0123548P.
XX
XX 23-MAR-1999; 99US-0125788P.
XX
XX 25-MAR-1999; 99US-0126264P.
XX
XX 29-MAR-1999; 99US-0126785P.
XX
XX 01-APR-1999; 99US-0127462P.
XX
XX 06-APR-1999; 99US-0128234P.
XX
XX 08-APR-1999; 99US-0128714P.
XX
XX 16-APR-1999; 99US-0129845P.
XX
XX 19-APR-1999; 99US-0130077P.
XX
XX 21-APR-1999; 99US-0130449P.
XX
XX 23-APR-1999; 99US-0130510P.
XX
XX 28-APR-1999; 99US-0130891P.
XX
XX 30-APR-1999; 99US-0132048P.
XX
XX 30-APR-1999; 99US-0132407P.
XX
XX 04-MAY-1999; 99US-0132484P.
XX
XX 05-MAY-1999; 99US-0132485P.
XX
XX 06-MAY-1999; 99US-0132486P.
XX
XX 06-MAY-1999; 99US-0132487P.
XX
XX 07-MAY-1999; 99US-0132863P.
XX
XX 11-MAY-1999; 99US-0134256P.
XX
XX 14-MAY-1999; 99US-0134218P.
XX
XX 14-MAY-1999; 99US-0134219P.
XX
XX 14-MAY-1999; 99US-0134370P.
XX
XX 18-MAY-1999; 99US-0134768P.
XX
XX 19-MAY-1999; 99US-0134941P.
XX
XX 20-MAY-1999; 99US-0135124P.
XX
XX 21-MAY-1999; 99US-0135353P.
XX
XX 24-MAY-1999; 99US-0135629P.
XX
XX 25-MAY-1999; 99US-0136021P.
XX
XX 27-MAY-1999; 99US-0136392P.
XX
XX 28-MAY-1999; 99US-0136782P.
XX
XX 01-JUN-1999; 99US-0137222P.
XX
XX 03-JUN-1999; 99US-0137528P.
XX
XX 04-JUN-1999; 99US-0137502P.
XX
XX 07-JUN-1999; 99US-0137724P.
XX
XX 08-JUN-1999; 99US-0138094P.

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148655P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 76.4%; Score 16.8; DB 3; Length 1359;
Best Local Similarity 90.0%; Pred No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCCATGATGGCGCATCTGAT 21

Db 830 CCCAGATGCCCACTGAT 849
|||||
RESULT 24
AAC36643
ID AAC36643 standard; DNA; 1363 BP.
XX
AC AAC36643;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14549.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 18-MAY-1999; 99US-0134370P.
XX 19-MAY-1999; 99US-0134768P.
XX 20-MAY-1999; 99US-0134941P.
XX 21-MAY-1999; 99US-0135124P.
XX 24-MAY-1999; 99US-0135353P.
XX 25-MAY-1999; 99US-0135629P.
XX 27-MAY-1999; 99US-0136021P.
XX 28-MAY-1999; 99US-0136392P.
XX 01-JUN-1999; 99US-0136782P.
XX 03-JUN-1999; 99US-0137222P.
XX 04-JUN-1999; 99US-0137528P.
XX 07-JUN-1999; 99US-0137502P.
XX 08-JUN-1999; 99US-0137724P.
XX 10-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 14-JUN-1999; 99US-0138847P.
XX 16-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 17-JUN-1999; 99US-0139453P.
XX 18-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148644P.
PR 16-AUG-1999; 99US-0149368P.

```
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 76.4%; Score 16.8; DB 3; Length 1363;
Best Local Similarity 90.0%; Pred. No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 2 CCCATGATGGCGCATCTGAT 21
DB 832 CCCAAGATGGCCCATCTGAT 851
```

RESULT 25
ABL05649
ID ABL05649 standard; cDNA; 1504 BP.

```
XX ABL05649;
AC 26-MAR-2002 (first entry)
DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 11429.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR P-PSDB; ABB61546.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX Claim 1; SEQ ID NO 11429; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1504 BP; 322 A; 404 C; 447 G; 331 T; 0 U; 0 Other;
Query Match 76.4%; Score 16.8; DB 4; Length 1504;
Best Local Similarity 90.0%; Pred. No. 3.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CCATGATGGCGCATCTGATG 22
DB 189 CAATGATGCCGATCTGATG 208
RESULT 26
ADT43185
ID ADT43185 standard; cDNA; 1629 BP.
XX AC ADT43185;
XX 02-DEC-2004 (first entry)
XX Bacterial polynucleotide #17936.
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
```

XX OS Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX PD
XX 20-FEB-2003; 2003US-00369493.
XX PF
XX 21-FEB-2002; 2002US-0360039P.
XX PR
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX PA
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX DR
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 41623; 122pp; English.
XX PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 1629 BP; 307 A; 579 C; 484 G; 259 T; 0 U; 0 Other;
SQ
Query Match 76.4%; Score 16.8; DB 13; Length 1629;
Best Local Similarity 90.0%; Pred. No. 3.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 CCCATGATCGCGCATCTGAT 21
Db 631 CCGATGATCGCGCATCTGAT 650
RESULT 27
ABL05648
ID ABL05648 standard; cDNA; 8018 BP.
XX AC
XX ABL05648;
XX 26-MAR-2002 (first entry)
XX DT
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 11426.
XX DE
XX

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX OS
XX Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEXE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-658860/75.
XX DR P-PSDB; ABB61545.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 11426; 21pp + Sequence Listing; English.
XX PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 8018 BP; 2333 A; 1646 C; 1725 G; 2314 T; 0 U; 0 Other;
SQ
Query Match 76.4%; Score 16.8; DB 4; Length 8018;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CCATGATCGCGCATCTGATG 22
Db 3201 CAATGATCGCGCATCTGATG 3220
RESULT 28
ABK12708/c
ID ABK12708 standard; DNA; 11197 BP.
XX AC
XX ABK12708;
XX 18-JUN-2002 (first entry)
XX DT
XX Human T1/ST2 receptor binding protein (IL1RL1LG) gene sequence.
XX DE
XX Human; T1/ST2 receptor binding protein; IL1RL1LG; cytostatic;
KW antiarthritic; gene therapy; isogene; single nucleotide polymorphism;
KW SNP; haplotyping; genotyping; cancer; arthritis; antibody; ASO;
KW allele-specific oligonucleotide; chromosome 19q13.2; gene; ds.
XX
XX Homo sapiens.
XX OS
XX Key Location/Qualifiers
FH variation replace(3867,A)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT /note= "polymorphic site 1, PSI"
FT variation replace(3868,A)
FT /*tag= b
FT


```
FT FT /*tag= f
FT FT /label= ORF6
FT FT /note= "Required for export of teicoplanin outside of the
FT FT cytoplasm, and for conferring resistance to teicoplanin
FT FT and required for regulating expression of one or more of
FT FT the genes of tcp gene cluster"
FT FT complement(7385..8065)
FT FT /*tag= g
FT FT /label= ORF7
FT FT /note= "Required for regulating expression of one or more
FT FT of the genes of tcp gene cluster"
FT FT 8924..10105
FT FT /*tag= h
FT FT /label= ORF8
FT FT /note= "Required for the addition and formation of N-acyl
FT FT -beta-glucosamine residues"
FT FT 10597..116827
FT FT /*tag= i
FT FT /label= ORF9
FT FT /note= "Required for synthesis of the heptapeptide
FT FT skeleton"
FT FT 16824..20009
FT FT /*tag= j
FT FT /label= ORF10
FT FT /note= "Required for synthesis of the heptapeptide
FT FT skeleton"
FT FT 20053..32256
FT FT /*tag= k
FT FT /label= ORF11
FT FT /note= "Required for synthesis of the heptapeptide
FT FT skeleton"
FT FT 32276..37873
FT FT /*tag= l
FT FT /label= ORF12
FT FT /note= "Required for synthesis of the heptapeptide
FT FT skeleton"
FT FT 37886..38095
FT FT /*tag= m
FT FT /label= ORF13
FT FT 38401..39222
FT FT /*tag= n
FT FT /label= ORF14
FT FT 39268..41049
FT FT /*tag= o
FT FT /label= ORF15
FT FT /note= "Required for attachment of mannosyl residue"
FT FT 41343..43289
FT FT /*tag= p
FT FT /label= ORF16
FT FT /note= "Required for export of teicoplanin outside of the
FT FT cytoplasm, and for conferring resistance to teicoplanin"
FT FT 43373..43585
FT FT /*tag= q
FT FT /label= ORF17
FT FT 43696..44871
FT FT /*tag= r
FT FT /label= ORF18
FT FT /note= "Required for cross-linking of aromatic residues
FT FT at position 2 and 4, 4 and 6, 1 and 3, and 5 and 7 of
FT FT teicoplanin"
FT FT 44894..46048
FT FT /*tag= s
FT FT /label= ORF19
FT FT /note= "Required for cross-linking of aromatic residues
FT FT at position 2 and 4, 4 and 6, 1 and 3, and 5 and 7 of
FT FT teicoplanin"
FT FT 46038..47234
FT FT /*tag= t
FT FT /label= ORF20
FT FT /note= "Required for cross-linking of aromatic residues
FT FT at position 2 and 4, 4 and 6, 1 and 3, and 5 and 7 of
FT FT teicoplanin"
FT FT 47484..49004
```

```
FT FT /*tag= u
FT FT /label= ORF21
FT FT /note= "Required for the chlorination of the aromatic
FT FT residues"
FT FT 49166..50344
FT FT /*tag= v
FT FT /label= ORF22
FT FT /note= "Required for cross-linking of aromatic residues
FT FT at position 2 and 4, 4 and 6, 1 and 3, and 5 and 7 of
FT FT teicoplanin"
FT FT 50461..51687
FT FT /*tag= w
FT FT /label= ORF23
FT FT /note= "Required for the addition and formation of N-acyl
FT FT -beta-glucosamine residues"
FT FT 51833..52804
FT FT /*tag= x
FT FT /label= ORF24
FT FT /note= "Required for the addition and formation of N-acyl
FT FT -beta-glucosamine residues"
FT FT 52948..54540
FT FT /*tag= y
FT FT /label= ORF25
FT FT /note= "Required for the beta-hydroxylation of the Tyr
FT FT residues"
FT FT 54757..56553
FT FT /*tag= z
FT FT /label= ORF26
FT FT /note= "Required for the addition and formation of N-acyl
FT FT -beta-glucosamine residues"
FT FT 56924..57979
FT FT /*tag= aa
FT FT /label= ORF27
FT FT 58746..59735
FT FT /*tag= ab
FT FT /label= ORF28
FT FT /note= "Required for regulating expression of one or more
FT FT of the genes of tcp gene cluster"
FT FT 60477..62855
FT FT /*tag= ac
FT FT /label= ORF29
FT FT /note= "Required for regulating expression of one or more
FT FT of the genes of tcp gene cluster"
FT FT 63036..64157
FT FT /*tag= ad
FT FT /label= ORF30
FT FT /note= "Required for synthesis of the 3,5-dihydroxy-
FT FT phenylglycine residues"
FT FT 64154..64819
FT FT /*tag= ae
FT FT /label= ORF31
FT FT /note= "Required for synthesis of the 3,5-dihydroxy-
FT FT phenylglycine residues"
FT FT 64816..66117
FT FT /*tag= af
FT FT /label= ORF32
FT FT /note= "Required for synthesis of the 3,5-dihydroxy-
FT FT phenylglycine residues"
FT FT 66146..66952
FT FT /*tag= ag
FT FT /label= ORF33
FT FT /note= "Required for synthesis of the 3,5-dihydroxy-
FT FT phenylglycine residues"
FT FT 66991..68352
FT FT /*tag= ah
FT FT /label= ORF34
FT FT /note= "Required for export of teicoplanin outside of the
FT FT cytoplasm, and for conferring resistance to teicoplanin"
FT FT 68349..68945
FT FT /*tag= ai
FT FT /label= ORF35
FT FT 69083..70180
FT FT /*tag= aj
```

```
FT /label= ORF36
FT /note= "Required for synthesis of the 4-hydroxy-
FT phenylglycine residues, and required for synthesis of the
FT 3,5-dihydroxy-phenylglycine residues"
FT CDS 70266..71321
FT /*tag= ak
FT /label= ORF37
FT /note= "Required for synthesis of the 4-hydroxy-
FT phenylglycine residues"
FT CDS 71318..72412

Query Match 76.4%; Score 16.8; DB 13; Length 73882;
Best Local Similarity 90.0%; Pred. NO. 6.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTGAT 21
Db 3888 CCCATGATGGCGCACCGAT 3869

RESULT 30
ABL67308
ID ABL67308 standard; DNA; 316 BP.
XX
AC ABL67308;
XX
XX 15-MAY-2002 (first entry)
XX
DE Thyroid cancer related gene sequence SEQ ID NO:5645.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
XX WO200194629-A2.
XX
XX 13-DEC-2001.
XX
XX 30-MAY-2001; 2001WO-US010838.
XX
XX 05-JUN-2000; 2000US-0209473P.
XX 05-JUN-2000; 2000US-0209531P.
XX 18-SEP-2000; 2000US-0233133P.
XX 18-SEP-2000; 2000US-0233617P.
XX 20-SEP-2000; 2000US-0234009P.
XX 20-SEP-2000; 2000US-0234034P.
XX 22-SEP-2000; 2000US-0234052P.
XX 22-SEP-2000; 2000US-0234509P.
XX 22-SEP-2000; 2000US-0234567P.
XX 25-SEP-2000; 2000US-0234923P.
XX 25-SEP-2000; 2000US-0234924P.
XX 25-SEP-2000; 2000US-0235077P.
XX 25-SEP-2000; 2000US-0235082P.
XX 25-SEP-2000; 2000US-0235134P.
XX 25-SEP-2000; 2000US-0235280P.
XX 26-SEP-2000; 2000US-0235637P.
XX 26-SEP-2000; 2000US-0235638P.
XX 27-SEP-2000; 2000US-0235711P.
XX 27-SEP-2000; 2000US-0235720P.
XX 27-SEP-2000; 2000US-0235840P.
XX 27-SEP-2000; 2000US-0235863P.
XX 28-SEP-2000; 2000US-0236028P.
XX 28-SEP-2000; 2000US-0236032P.
XX 28-SEP-2000; 2000US-0236033P.
XX 28-SEP-2000; 2000US-0236034P.
XX 28-SEP-2000; 2000US-0236109P.
XX 28-SEP-2000; 2000US-0236111P.
XX 29-SEP-2000; 2000US-0236842P.
XX 29-SEP-2000; 2000US-0236891P.
XX 02-OCT-2000; 2000US-0237172P.

PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237318P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
XX Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 5645; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of M1, and the data is sufficient to convey the chemical structure and/or
XX properties of the agent. M1 can be used in the treatment of cancer such
XX as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
XX prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
XX tumour
XX
XX Sequence 316 BP; 77 A; 67 C; 85 G; 87 T; 0 U; 0 Other;
XX
XX Query Match 74.5%; Score 16.4; DB 6; Length 316;
XX Best Local Similarity 94.4%; Pred. NO. 4e+02;
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTG 19
Db 178 CCCATGATGGCACATCTG 195

RESULT 31
ABN94178
ID ABN94178 standard; DNA; 316 BP.
XX
XX AC ABN94178;
XX
XX 13-AUG-2002 (first entry)
XX
XX Gene #676 used to diagnose liver cancer.
XX
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
XX metastatic liver tumour; cytostatic; expression profile; disease state;
XX disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
XX Homo sapiens.
XX
XX WO200229103-A2.
XX
```



```

PR 06-APR-2000; 2000US-0194941P.
XX (GEMY ) GENETICS INST INC.
XX
XX Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
XX Gulukota K, Graham JR;
XX
XX WPI; 2002-179323/23.
XX
XX Six hundred and twenty five polynucleotides derived from a variety of
XX human tissue sources which encode secreted proteins, useful for treating
XX immune deficiencies and disorders such as autoimmune disorders.
XX
XX Claim 1; Page 105; 339pp; English.
XX
XX The invention relates to 625 polynucleotides which have been derived from
XX a variety of human tissue sources and which encode novel secreted
XX proteins, their complements and sequences that hybridise to them. Also
XX included are a vector comprising the polynucleotide, a host cell
XX transformed with the vector, the proteins encoded by the polynucleotides,
XX antibodies that bind to the proteins and identification of modulators of
XX the proteins or the expression of the polynucleotide. The polynucleotides
XX can be used as probes for the identification and isolation of full length
XX cDNA and genomic DNA. The polynucleotides and proteins can also be used
XX as nutritional supplements. The protein is useful in the treatment of
XX various immune deficiencies and disorders such as viral infections,
XX bacterial infections, fungal infections, autoimmune disorders (e.g.
XX rheumatoid arthritis, multiple sclerosis, autoimmune thyroiditis and
XX diabetes) and allergic reactions and conditions (e.g. asthma). They are
XX also useful for treating neurodegenerative diseases (e.g. Alzheimer's
XX disease, Parkinson's disease), liver fibrosis, coagulation disorders
XX (e.g. haemophilia), inflammatory disorders (e.g. Crohn's disease) and
XX tumours. They are also useful for tissue regeneration, for wound healing
XX and in the treatment of burns, incisions and ulcers. The proteins are
XX also useful for regulating haematopoiesis, for treating myeloid or
XX lymphoid cell deficiencies. The present sequence is one of the 625 cDNA
XX sequences encoding a secreted protein
XX
XX Sequence 600 BP; 104 A; 174 C; 158 G; 164 T; 0 U; 0 Other;
XX
XX Query Match 74.5%; Score 16.4; DB 6; Length 600;
XX Best Local Similarity 94.4%; Pred. No. 4.5e+02;
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 2 CCCATGATGGCGCATCTG 19
XX |||||
XX Db 456 CCCATGATGGCGCATCTG 439
XX
XX RESULT 34
XX ABK73908
XX ID ABK73908 standard; DNA; 852 BP.
XX
XX AC ABK73908;
XX
XX DT 13-AUG-2002 (first entry)
XX
XX DE Bacillus licheniformis genomic sequence tag (GST) #1199.
XX
XX KW Differential gene expression; genomic sequenced tag; GST;
XX KW altered culture condition; environmental stress;
XX KW physiological provocation; ds.
XX
XX OS Bacillus licheniformis.
XX
XX PN WO200229113-A2.
XX
XX PD 11-APR-2002.
XX
XX PF 05-OCT-2001; 2001WO-US031437.
XX
XX PR 06-OCT-2000; 2000US-00680598.
XX 27-MAR-2001; 2001US-0279526F.
XX
XX (NOVO ) NOVOZYMES BIOTECH INC.
XX (NOVO ) NOVOZYMES AS.
XX
XX Berka R, Clausen IG;
XX
XX WPI; 2002-416684/44.
XX
XX Monitoring differential expression of several genes in first Bacillus
XX cell relative to expression of same genes in one or more second Bacillus
XX cells, by using substrate containing Bacillus genomic sequenced tag
XX array.
XX
XX Claim 4; SEQ ID NO 1199; 200pp; English.
XX
XX The invention describes a method of monitoring differential expression of
XX genes in a first Bacillus cell relative to expression of the genes in
XX other Bacillus cells, comprising hybridising labelled nucleic acid probes
XX isolated from Bacillus cells to a substrate containing array of Bacillus
XX genomic sequenced tags (GST), examining the array, and determining
XX relative gene expression by an observed hybridisation reporter signal of
XX a spot in the array. The method is useful for measuring the expression of
XX genes in a first Bacillus cell relative to expression of the same genes
XX in one or more second Bacillus cells. The method is useful for monitoring
XX global expression of several genes from a Bacillus cell, discovering new
XX genes, identifying possible functions of unknown open reading frames and
XX monitoring gene copy number variation and stability. Monitoring changes
XX in expression of genes may be used to provide a representation of the way
XX in which Bacillus cells adapt to changes in culture conditions,
XX environmental stress or other physiological provocation. Extensive follow
XX up characterisation is unnecessary, when one spot on an array equals one
XX gene or one open reading frame, since sequence information is available.
XX This sequence represents a genomic sequence tag (GST) used in the method
XX of the invention. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 852 BP; 214 A; 194 C; 238 G; 205 T; 0 U; 1 Other;
XX
XX Query Match 74.5%; Score 16.4; DB 6; Length 852;
XX Best Local Similarity 94.4%; Pred. No. 4.8e+02;
XX Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 5 ATGATGGCGCATCTGATG 22
XX |||||
XX Db 363 ATGATGGCGCATCTGATG 380
XX
XX RESULT 35
XX AAF18110/c
XX ID AAF18110 standard; DNA; 1266 BP.
XX
XX AC AAF18110;
XX
XX DT 14-MAR-2001 (first entry)
XX
XX DE Lung cancer associated polynucleotide sequence SEQ ID 129.
XX
XX KW Human; lung cancer associated protein; neuroprotective; cytostatic;
XX KW cardioactive; immunomodulatory; muscular active; vulnerary;
XX KW gastrointestinal; nephrotropic; antiinfective; gynecological;
XX KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
XX KW proliferative disorder; wound healing; infectious disease; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200055180-A2.
XX
XX PD 21-SEP-2000.
XX
XX PF 08-MAR-2000; 2000WO-US005918.
XX 12-MAR-1999; 99US-0124270P.
XX

```

```
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX Ruben SM;
XX
XX WPI; 2000-587514/55.
DR P-PSDB; AAB56234.
XX
XX Lung cancer associated gene sequences, referred to as lung cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as lung cancer.
XX
XX Claim 1; Page 597; 1425pp; English.
XX
XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
CC associated proteins and polynucleotide sequences, their agonists, and
CC antagonists may have neuroprotective; cytostatic; cardioactive;
CC immunomodulatory; muscular active general; vulnery; gastrointestinal
CC general; nephrotropic; antiinfective; gynecological; or antibacterial
CC activity. The invention also includes antibodies specific for the protein
CC or polynucleotide sequences. The lung cancer associated polynucleotide
CC sequences may be used for detection of lung cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The proteins may be used to treat disorders such as
CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
CC cardiovascular, renal, and proliferative disorders. The proteins may also
CC be used in the treatment of wounds and infectious diseases.
CC Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
CC used in the course of the invention for the identification and
CC characterisation of the polynucleotide and protein sequences
XX
XX Sequence 1266 BP; 264 A; 377 C; 311 G; 307 T; 0 U; 7 Other;
SQ
Query Match 74.5%; Score 16.4; DB 3; Length 1266;
Best Local Similarity 94.4%; Pred. No. 5.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CCCATGATGGGCATCTG 19
Db 961 CCCATGATGGGCATCTG 944
RESULT 36
AAH98188/c
ID AAH98188 standard; cDNA; 1712 BP.
AC AAH98188;
XX
XX 12-OCT-2001 (first entry)
XX
XX Human EST-derived coding sequence SEQ ID NO: 45.
DE
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; Gene mapping; genetic disorder; biodiversity;
KW gene therapy; nutrition; ss.
XX
XX Homo sapiens.
OS
XX WO200154477-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US002687.
XX
XX 25-JAN-2000; 2000US-00491404.
PR 17-JUL-2000; 2000US-00617746.
PR 03-AUG-2000; 2000US-00631451.
PR 15-SEP-2000; 2000US-00663870.
XX
XX (HYSE-) HYSEQ INC.
PA
```

```
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werthman T;
XX
XX WPI; 2001-476164/51.
DR P-PSDB; AAM23529.
XX
XX Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use.
PT
XX Claim 1; Page 212-213; 1275pp; English.
XX
XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention
XX
XX Sequence 1712 BP; 309 A; 579 C; 440 G; 382 T; 0 U; 2 Other;
SQ
Query Match 74.5%; Score 16.4; DB 4; Length 1712;
Best Local Similarity 94.4%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 CCCATGATGGGCATCTG 19
Db 1556 CCCATGATGGGCATCTG 1539
RESULT 37
AAZ43799/c
ID AAZ43799 standard; cDNA; 1773 BP.
XX
XX AAZ43799;
AC
XX 10-MAR-2000 (first entry)
XX
XX Human fetal brain cDNA clone vc30_1.
XX
XX Human; secreted protein; treatment; nutritional activity; cytokine;
KW cell proliferation; cell differentiation; hematopoiesis regulation;
KW tissue growth; activin; inhibitor; chemotactic; chemokinetic; hemostatic;
KW thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
KW gene therapy; ds.
XX
XX Homo sapiens.
OS
XX WO9955721-A1.
XX
XX 04-NOV-1999.
XX
XX 23-APR-1999; 99WO-US008504.
XX
XX 24-APR-1998; 98US-0082904P.
PR 11-JUN-1998; 98US-0088994P.
PR 12-JUN-1998; 98US-0089278P.
PR 02-JUL-1998; 98US-0091647P.
PR 24-AUG-1998; 98US-0097639P.
PR 22-APR-1999; 99US-00097639.
XX
XX (ALPH-) ALPHAGENE INC.
XX
XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX
XX WPI; 2000-052801/04.
DR P-PSDB; AAY50934, AAY50935.
XX
XX New polynucleotides encoding secreted human proteins, derived from human
PT fetal brain, adult skin, adult brain, adult heart, adult thymus and adult
PT aorta cDNA libraries.
PA
```

XX Claim 54a; Page 248; 282pp; English.

XX This invention describes novel human secreted proteins which are encoded

CC by polynucleotides obtained from fetal brain, adult skin, adult brain,

CC adult heart, adult thymus and adult aorta cDNA libraries. The

CC polynucleotides and proteins are predicted to have biological activities

CC which would make them suitable for treating, preventing or ameliorating

CC medical conditions in humans and animals, although no supporting data is

CC given. Suggested activities include nutritional activity, cytokine and

CC cell proliferation/differentiation activity, immune stimulating (e.g. as

CC vaccines) or suppressing activity, hematopoiesis regulating activity,

CC tissue growth activity, activin/inhibin activity,

CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,

CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor

CC invasion suppressor activity, and tumor inhibition activity. The

CC polynucleotides are also stated to be useful for gene therapy. AAZ43777-

CC 243808 represent the polynucleotides described in the invention which

CC encode the proteins represented in AAY50905-Y50947

XX

SQ Sequence 1773 BP; 324 A; 602 C; 459 G; 388 T; 0 U; 0 Other;

Query Match 74.5%; Score 16.4; DB 3; Length 1773;

Best Local Similarity 94.4%; Pred. No. 5.4e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGGCATCTG 19

|||||

Db 1605 CCCATGATGGGCATCTG 1598

RESULT 38

ABL55085/c

ID ABL55085 standard; DNA; 1984 BP.

AC ABL55085;

XX

XX 08-OCT-2002 (first entry)

XX

DE Human cDNA sequence #1 from clone HLCND09.

XX

KW Human; HLCND09; secreted protein; immunosuppressive; food preservative;

KW antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;

KW vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;

KW virucide; fungicide; ophthalmological; vulnery; gene therapy; ELISA;

KW radioimmunoassay; enzyme linked immunosorbent assay; autoimmune disease;

KW hyperproliferative disorder; cardiovascular disorder; angiogenesis;

KW cerebrovascular disorder; nervous system disorder; ocular disorder;

KW wound healing; food additive; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 146..478

FT /*tag= a

FT sig_peptide 146..259

FT /*tag= b

FT mat_peptide 260..475

FT /*tag= c

XX

FN WO200222638-A1.

XX

XX 21-MAR-2002.

XX

XX 17-JAN-2001; 2001WO-US001386.

XX

XX 12-SEP-2000; 2000US-0232104P.

XX

XX (HUMA-) HUMAN GENOME SCI INC.

XX

XX Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR;

PI Olsen HS, Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH;

PI Fiscella M, Ni J;

XX

XX WPI; 2002-258041/30.

DR P-PSDB; ABB77016.

XX

XX New nucleic acid molecules encoding 22 human secreted proteins for

PT diagnosing or treating e.g. autoimmune diseases, hyperproliferative

PT disorders, and cardiovascular disorders, and used as food additives or

PT preservatives.

XX

PS Disclosure; Page 449-450; 526pp; English.

XX

XX The sequence represents a cDNA sequence of the invention, isolated from

CC human clone ID HLCND09. The invention relates to novel isolated nucleic

CC acid molecules encoding 22 human secreted proteins. The proteins of the

CC invention have immunosuppressive, antiarthritic, antirheumatic,

CC antiproliferative, cytostatic, cardiant, vasotropic, cerebroprotective,

CC neurotropic, neuroprotective, antibacterial, virucide, fungicide,

CC ophthalmological, and vulnery activity. The polynucleotides may have a

CC use in gene therapy. The polynucleotides and polypeptides encoded by them

CC are used to prevent, treat or ameliorate a medical condition in e.g.

CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. The

CC polynucleotides and polypeptides are also used in diagnosing a

CC pathological condition or susceptibility to a pathological condition. The

CC antibodies to the proteins can also be used in alleviating symptoms

CC associated with the disorders and in diagnostic immunoassays e.g.

CC radioimmunoassays or enzyme linked immunosorbent assays (ELISA).

CC Disorders which are diagnosed or treated include autoimmune diseases,

CC hyperproliferative disorders, cardiovascular disorders, cerebrovascular

CC disorders, angiogenesis, nervous system disorders, infections caused by

CC bacteria, viruses and fungi and ocular disorders. The polypeptides can

CC also be used to aid wound healing and epithelial cell proliferation. The

CC polypeptides can also be used as a food additive or preservative

XX

SQ Sequence 1984 BP; 385 A; 638 C; 524 G; 435 T; 0 U; 2 Other;

Query Match 74.5%; Score 16.4; DB 6; Length 1984;

Best Local Similarity 94.4%; Pred. No. 5.6e+02;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGGCATCTG 19

|||||

Db 1789 CCCATGATGGGCATCTG 1772

RESULT 39

ADA39911/c

ID ADA39911 standard; cDNA; 1984 BP.

XX

AC ADA39911;

XX

XX 20-NOV-2003 (first entry)

XX

XX Human secreted protein encoding cDNA.

XX

XX Human; secreted protein; cancer; hyperproliferative disorder;

XX rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;

XX anaemia; allergic reaction; asthma; cardiovascular disorder;

XX wound healing; cytostatic; immunosuppressive; neurotropic; neuroprotective;

XX antiviral; anti-allergic; hepatotropic; antidiabetic; anti-inflammatory;

XX vulnery; cardiant; gene therapy; ss.

XX

OS Homo sapiens.

XX

XX WO2002102993-A2.

XX

XX 27-DEC-2002.

XX

XX 19-MAR-2002; 2002WO-US008123.

XX

XX 21-MAR-2001; 2001US-0277340P.

PR 19-JUL-2001; 2001US-0306171P.

PR 13-NOV-2001; 2001US-0331287P.

XX

PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Rosen CA, Ruben SM;
XX	
DR	WPT; 2003-175238/17.
XX	
PT	New human secreted proteins and nucleic acid molecules, useful for
PT	preparing a diagnostic or pharmaceutical composition for diagnosing,
PT	preventing or treating cancer or other hyperproliferative disorder,
PT	asthma, allergies or AIDS.
XX	
PS	Claim 9; SEQ ID NO 293; 3205pp; English.
XX	
CC	The invention relates to novel genes AD39629-ADA40565 and proteins
CC	ADA40566-ADA41501 for human secreted proteins, useful for preventing,
CC	treating or ameliorating medical conditions e.g. by protein or gene
CC	therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC	fragments, and agonists or antagonists that bind to the polypeptide are
CC	useful for preparing a diagnostic or pharmaceutical composition for
CC	diagnosing or treating cancer or other hyperproliferative disorder. The
CC	polypeptides and nucleic acid molecules are also useful for detecting,
CC	preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC	or other hyperproliferative disorders including neoplasms, autoimmune
CC	disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC	erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC	anaemia), haematopoietic or haematological disorders (e.g. anaemia,
CC	thrombocytopenia), allergic reactions including asthma or eczema,
CC	inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC	bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC	Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC	(e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC	fungal or viral infections including HIV/AIDS), or wound healing and
CC	disorders of epithelial cell proliferation. The nucleic acids are also
CC	useful for chromosome identification, radiation hybrid mapping or long-
CC	range restriction mapping, as molecular weight markers, or as
CC	hybridization or diagnostic probes. The polypeptides and antibodies are
CC	useful for providing immunologic probes for differential identification
CC	of the tissues immunohistochemistry assays. Note: The sequence data for
CC	this patent did not form part of the printed specification, but was
CC	obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 1984 BP; 385 A; 638 C; 524 G; 435 T; 0 U; 2 Other;
	Query Match 74.5%; Score 16.4; DB 8; Length 1984;
	Best Local Similarity 94.4%; Pred. No. 5.6e+02;
	Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	2 CCCATGATGGCGCATCTG 19 Db 1789 CCCATGATGGCATCTG 1772
RESULT 40	
ADAS6100/c	
ID ADAS6100 standard; DNA; 1984 BP.	
XX AC ADAS6100;	
XX	
DT 20-NOV-2003 (first entry)	
XX	
DE Gene encoding human secreted protein #279.	
XX	
KW immunosuppressive; antinflammatory; antiasthmatic; antiallergic;	
KW cycostatic; cerebroprotective; neuroprotective; nootropic;	
KW cardiovascular; antiarteriosclerotic; gene therapy;	
KW human secreted protein; immune disorder; inflammation;	
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;	
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;	
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;	
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;	
KW triple helix formation; antisense gene therapy; forensic biology; ds;	
KW gene.	

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:48:59 ; Search time 48.9467 Seconds
(without alignments)
735.454 Million cell updates/sec

Title: US-09-674-277-24

Perfect score: 22

Sequence: 1 accatgatggcgatctgatg 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/5A-COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B-COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A-COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B-COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS-COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.2	78.2	672	4	US-09-489-039A-907
2	16.8	76.4	312	4	US-09-489-039A-2742
3	16.8	76.4	759	4	US-09-489-039A-3002
4	16.8	76.4	7845	4	US-09-949-016-14467
5	16.2	73.6	351	4	US-09-248-796A-11892
6	16.2	73.6	732	4	US-09-602-777A-325
7	16.2	73.6	1183	4	US-09-270-767-14467
8	16.2	73.6	1860	4	US-09-934-903-1
9	16.2	73.6	1860	4	US-09-934-868-61
10	16.2	73.6	5430	3	US-09-012-515A-11
11	16.2	73.6	5430	3	US-08-360-144A-11
12	16.2	73.6	5430	3	US-09-012-504A-11
13	16.2	73.6	5430	4	US-09-012-399A-11
14	16.2	73.6	7111	4	US-09-949-016-2128
15	16.2	73.6	7653	3	US-08-471-112A-1
16	16.2	73.6	7653	4	US-09-950-634-1
17	16.2	73.6	7824	5	PCI-US95-06722-11
18	16.2	73.6	11575	4	US-09-938-956-5
19	16.2	73.6	155266	4	US-09-949-016-13870
20	16.2	73.6	238815	4	US-09-949-016-16274
21	15.8	71.8	1140	4	US-09-107-532A-2426
22	15.8	71.8	36519	3	US-08-923-137-2
23	15.6	70.9	399	4	US-09-540-236-764
24	15.6	70.9	708	4	US-09-328-352-3714
25	15.6	70.9	938	4	US-09-270-767-16924
26	15.6	70.9	938	4	US-09-270-767-16924
27	15.6	70.9	956	4	US-09-949-016-1403

28	15.6	70.9	1080	1	US-08-225-757B-1	Sequence 1, Appli
29	15.6	70.9	1080	2	US-08-722-050-1	Sequence 1, Appli
30	15.6	70.9	1080	4	US-09-883-985-1	Sequence 1, Appli
31	15.6	70.9	1236	4	US-09-540-236-631	Sequence 631, App
32	15.6	70.9	1292	4	US-09-270-767-11046	Sequence 11046, A
33	15.6	70.9	1650	4	US-09-461-325-75	Sequence 75, Appl
34	15.6	70.9	1650	4	US-10-012-542-75	Sequence 75, Appl
35	15.6	70.9	1650	4	US-10-115-123-75	Sequence 75, Appl
36	15.6	70.9	1939	4	US-09-949-016-3778	Sequence 3778, Ap
37	15.6	70.9	1962	4	US-09-540-236-425	Sequence 425, App
38	15.6	70.9	2612	4	US-09-495-050A-214	Sequence 214, App
39	15.6	70.9	5738	1	US-08-409-995-3	Sequence 3, Appli
40	15.6	70.9	5738	3	US-08-685-467-3	Sequence 3, Appli
41	15.6	70.9	7253	3	US-09-268-347-35	Sequence 35, Appl
42	15.6	70.9	7291	3	US-08-913-942-3	Sequence 3, Appli
43	15.6	70.9	7291	4	US-09-684-707-3	Sequence 3, Appli
44	15.6	70.9	17527	4	US-09-949-016-13145	Sequence 13145, A
45	15.6	70.9	29172	4	US-09-949-016-15520	Sequence 15520, A

ALIGNMENTS

RESULT 1

US-09-489-039A-907

; Sequence 907, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 907

; LENGTH: 672

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-907

Query Match 78.2%; Score 17.2; DB 4; Length 672;
Best Local Similarity 86.4%; Pred. No. 56;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGCGCATCTGATG 22

Db 300 ACCGATGCTGGCGAGCTGATG 321

RESULT 2

US-09-489-039A-2742/c

; Sequence 2742, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 2742

; LENGTH: 312

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-2742

Query Match 76.4%; Score 16.8; DB 4; Length 312;

Best Local Similarity 90.0%; Pred. No. 77;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCATGATGGCGATCTGATG 22
||||| ||||| ||||| ||||| |||||
Db 30 CCATGCTGGCGGATCTGATG 11

RESULT 3

US-09-489-039A-3002
; Sequence 3002, Application US/09489039A
; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 3002

; LENGTH: 759

; TYPE: DNA

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-3002

Query Match 76.4%; Score 16.8; DB 4; Length 759;

Best Local Similarity 90.0%; Pred. No. 90;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCATGATGGCGATCTGATG 22
||||| ||||| ||||| ||||| |||||
Db 545 CCATGCTGGCGGATCTGATG 564

RESULT 4

US-09-949-016-14467/c

; Sequence 14467, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 14467

; LENGTH: 7845

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-14467

Query Match 76.4%; Score 16.8; DB 4; Length 7845;

Best Local Similarity 90.0%; Pred. No. 1.4e+02;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 CCATGATGGCGATCTGATG 22
||||| ||||| ||||| ||||| |||||
Db 1819 CCATGATGGCGATCTTTTG 1800

RESULT 5

US-09-248-796A-11892/c

; Sequence 11892, Application US/09248796A

; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 11892

; LENGTH: 351

; TYPE: DNA

; ORGANISM: Candida albicans

US-09-248-796A-11892

Query Match 73.6%; Score 16.2; DB 4; Length 351;

Best Local Similarity 85.7%; Pred. No. 1.6e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 278 ACCCATGATGGCGTGTGAT 258

RESULT 6

US-09-602-777A-325/c

; Sequence 325, Application US/09602777A

; Patent No. 6831165

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Markus

; APPLICANT: Kroger, Burkhard

; APPLICANT: Schroder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Haberhauser, Gregor

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS

; FILE REFERENCE: INVOLVED IN HOMEOSTASIS AND ADAPTATION

; CURRENT APPLICATION NUMBER: US/09/602,777A

; CURRENT FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: US 60/141031

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: DE 19931636.8

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19932125.6

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932126.4

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932127.2

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932128.0

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932129.9

; PRIOR FILING DATE: 1999-07-19

; PRIOR APPLICATION NUMBER: DE 19932226.0

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932920.6

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19932922.2

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19932924.9

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19932928.1

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19932930.3

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19932933.8

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19932935.4

RESULT 8
US-09-934-903-1
; Sequence 1, Application US/09934903
; Patent No. 6660507
; GENERAL INFORMATION:

RESULT 10
US-09-012-515A-11
; Sequence 11, Application US/09012515A
; Patent No. 6127521
; GENERAL INFORMATION:
; APPLICANT: Berlin, Vivian
; APPLICANT: Chiu, Maria Isabel

APPLICANT: Cottarel, Guillaume
APPLICANT: Damagnez, Veronique
TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,515A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/360,144
FILING DATE: 20-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 5430 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..5427
US-09-012-515A-11
Query Match 73.6%; Score 16.2; DB 3; Length 5430;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCCATGATGGGCATCTGAT 21
Db 2398 ACCCATGATGGGCATTTTAT 2418
RESULT 11
US-08-360-144A-11
; Sequence 11, Application US/08360144A
; Patent No. 6150137
; GENERAL INFORMATION:
; APPLICANT: Berlin, Vivian
; APPLICANT: Chiu, Maria Isabel
; APPLICANT: Cottarel, Guillaume
; APPLICANT: Damagnez, Veronique
; TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/360,144A
FILING DATE: 20-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: APV-036.02
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 5430 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..5427
US-08-360-144A-11
Query Match 73.6%; Score 16.2; DB 3; Length 5430;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCCATGATGGGCATCTGAT 21
Db 2398 ACCCATGATGGGCATTTTAT 2418
RESULT 12
US-09-012-504A-11
; Sequence 11, Application US/09012504A
; Patent No. 6464974
; GENERAL INFORMATION:
; APPLICANT: Berlin, V.
; APPLICANT: Chiu, I.
; APPLICANT: Cottarel, G.
; APPLICANT: Damagnez, V.
; TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
; FILE REFERENCE: APBI-P05-036
; CURRENT APPLICATION NUMBER: US/09/012,504A
; CURRENT FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 08/360,144
; PRIOR FILING DATE: 1994-12-20
; PRIOR APPLICATION NUMBER: 08/250,795
; PRIOR FILING DATE: 1994-05-27
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 5430
; TYPE: DNA
; ORGANISM: Mammalian
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(5427)
; OTHER INFORMATION:
US-09-012-504A-11
Query Match 73.6%; Score 16.2; DB 3; Length 5430;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCCATGATGGGCATCTGAT 21
Db 2398 ACCCATGATGGGCATTTTAT 2418
RESULT 13
US-09-012-399A-11

; Sequence 11, Application US/09012399A
; Patent No. 6509152
; GENERAL INFORMATION:
; APPLICANT: Berlin, Vivian
; APPLICANT: Chiu, Maria Isabel
; APPLICANT: Cottarel, Guillaume
; APPLICANT: Danaghez, Veronique
; TITLE OF INVENTION: IMMUNOSUPPRESSANT TARGET PROTEINS
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/012,399A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/360,144
; FILING DATE: 20-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: APV-036.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5430 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..5427
; US-09-012-399A-11

Query Match 73.6%; Score 16.2; DB 4; Length 5430;
Best Local Similarity 85.7%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGAT 21
Db 2398 ACCCATGATGGGCATTTAT 2418

RESULT 14
US-09-949-016-2128
; Sequence 2128, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 2128
; LENGTH: 7111
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2128

Query Match 73.6%; Score 16.2; DB 4; Length 7111;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGAT 21
Db 4451 ACCCATGATGGGCATTTAT 4471

RESULT 15
US-08-471-112A-1
; Sequence 1, Application US/08471112A
; Patent No. 6313284
; GENERAL INFORMATION:
; APPLICANT: Molnar-Kimber, Katherine L.
; APPLICANT: Failli, Amedeo F.
; APPLICANT: Caggiano, Thomas J.
; APPLICANT: Nakanishi, Koji
; APPLICANT: Chen, Yanqiu
; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,112A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/384,524
; FILING DATE: 13-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/312,023
; FILING DATE: 26-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,975
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Siekman, Michael T.
; REGISTRATION NUMBER: 36,276
; REFERENCE/DOCKET NUMBER: 01142.0058-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7653 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; US-08-471-112A-1

Query Match 73.6%; Score 16.2; DB 3; Length 7653;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGGCATCTGAT 21
|||||
Db 4621 ACCCATGATGGGCATTTAT 4641

RESULT 16

US-09-950-634-1
; Sequence 1, Application US/09950634
; Patent No. 6713607

GENERAL INFORMATION:

APPLICANT: Molnar-Kimber, Katherine L.
Failli, Amedeo F.
Caggiano, Thomas J.
Nakanishi, Koji
Chen, Yangiu

TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSER: Finnegan, Henderson, Farabow, Garrett &
Dunner, L.L.P.

STREET: 1300 I Street, N.W.

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

APPLICATION DATA:

APPLICATION NUMBER: US/09/950,634

FILING DATE: 13-Sep-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/471,112

FILING DATE: 06-JUN-1995

APPLICATION NUMBER: US 08/384,524

FILING DATE: 13-FEB-1995

APPLICATION NUMBER: US 08/312,023

FILING DATE: 26-SEP-1995

APPLICATION NUMBER: US 08/207,975

FILING DATE: 08-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Siekman, Michael T.

REGISTRATION NUMBER: 36,276

REFERENCE/DOCKET NUMBER: 01142.0058-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-408-4000

TELEFAX: 202-408-4400

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 7653 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-950-634-1

Query Match 73.6%; Score 16.2; DB 4; Length 7653;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGGCATCTGAT 21

|||||
Db 4621 ACCCATGATGGGCATTTAT 4641

RESULT 17

PCT-US95-06722-11

; Sequence 11, Application PC/TUS9506722
; GENERAL INFORMATION:

APPLICANT:

TITLE OF INVENTION: Immunosuppressant Target Proteins

NUMBER OF SEQUENCES: 25

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII (text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/06722

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/250,795

FILING DATE: 27-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/250,795

FILING DATE: 20-DEC-1994

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 7824 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 97..7743

PCT-US95-06722-11

Query Match 73.6%; Score 16.2; DB 5; Length 7824;
Best Local Similarity 85.7%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGGCATCTGAT 21

|||||
Db 4714 ACCCATGATGGGCATTTAT 4734

RESULT 18

US-09-938-956-5

; Sequence 5, Application US/09938956

; Patent No. 6818424

GENERAL INFORMATION:

APPLICANT: Wang, Siqun

APPLICANT: Dicosimo, Deana J.

APPLICANT: Koffas, Mattheos

APPLICANT: Odom, J. Martin

TITLE OF INVENTION: Production of Monoterpene

FILE REFERENCE: CIL809 US NA

CURRENT APPLICATION NUMBER: US/09/938,956

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: 60/229,907

PRIOR FILING DATE: 2000-09-0

PRIOR APPLICATION NUMBER: 60/229,858

PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 7

SOFTWARE: Microsoft Office 97

SEQ ID NO 5

LENGTH: 11575

TYPE: DNA

ORGANISM: Plasmid

US-09-938-956-5

Query Match 73.6%; Score 16.2; DB 4; Length 11575;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGGCATCTGAT 21

|||||
Db 2028 ACCCATGCTGGCGCTTTGAT 2048

RESULT 19
US-09-949-016-13870
; Sequence 13870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13870
; LENGTH: 155266
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13870

Query Match 73.6%; Score 16.2; DB 4; Length 155266;
Best Local Similarity 85.7%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
DB 114677 ACCCATGATGGCGCATTTAT 114697

RESULT 20
US-09-949-016-16274
; Sequence 16274, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16274
; LENGTH: 238815
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16274

Query Match 73.6%; Score 16.2; DB 4; Length 238815;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTGATG 22
DB 97929 CCAGGATGGCGAATATGATG 97949

RESULT 21
US-09-107-532A-2426
; Sequence 2426, Application US/09107532A

; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denise
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 2426:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1140 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1140
; SEQUENCE DESCRIPTION: SEQ ID NO: 2426:
US-09-107-532A-2426
Query Match 71.8%; Score 15.8; DB 4; Length 1140;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 CATGATGGCGCATCTGATG 22
DB 511 CAAGATGGCGCATCGATG 529
RESULT 22
US-08-923-137-2
; Sequence 2, Application US/08923137
; Patent No. 6083716
; GENERAL INFORMATION:
; APPLICANT: Wilson, James M.
; APPLICANT: Farina, Steven P.
; APPLICANT: Fisher, Krishna J.
; TITLE OF INVENTION: Chimpanzee Adenovirus Vectors
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr., P.O. Box 457
; CITY: Spring House

STATE: Pennsylvania
COUNTRY: United States of America
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/923,137
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/024,700
FILING DATE: 06-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: GNPVN.021CIPUSA
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-540-9200
TELEFAX: 215-540-5818
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 36519 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-923-137-2

```

Query Match          71.8%; Score 15.8; DB 3; Length 36519;
Best Local Similarity 89.5%; Pred. No. 5.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 2 CCATGATGGCGCATCTGA 20
|||
Db 27129 CCCATGGTGGCGCAGCTGA 27147

```

RESULT 23
US-09-540-236-764
; Sequence 764, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 764
; LENGTH: 399
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-764

```

```
Query Match          70.9%; Score 15.6; DB 4; Length 399;
Best Local Similarity 81.8%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy 1 ACCCATGATGGCGCATCTGATG 22
||| ||| ||| ||| ||| ||| ||| |||
Db 265 ACCGATGATGGCGTAAATGATG 286

RESULT 24
US-09-328-352-3714
; Sequence 3714, Application US/09328352
; Patent No. 6562358
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.

```

: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
:
: TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
:
: FILE REFERENCE: GTC99-03PA
:
: CURRENT APPLICATION NUMBER: US/09/328.352
:
: CURRENT FILING DATE: 1999-06-04
:
: NUMBER OF SEQ ID NOS: 8252
:
: SEQ ID NO 3714
:
: LENGTH: 708
:
: TYPE: DNA
:
: ORGANISM: Acinetobacter baumannii
:
: US-09-328-352-3714

```

```
Query Match          70.9%; Score 15.6; DB 4; Length 708;
Best Local Similarity 81.8%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy 1 ACCCATGATGGCGCATCTGATG 22
133 ACTCATGAAGGCCCATCTGTTG 154
Db

```

RESULT 25
US-09-270-767-1642
; Sequence 1642, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1642
; LENGTH: 938
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1642

```

Query Match 70.9%; Score 15.6; DB 4; Length 938;
Best Local Similarity 81.8%; Pred. No. 3.7e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGCGCATCTGATG 22
 ||| ||||| |
Db 76 ACCACTGATGGCTCAACTGATG 97

```

RESULT 26
US-09-270-767-16924
; Sequence 16924, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16924
; LENGTH: 938
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-16924

```

```
Query Match      70.9%; Score 15.6; DB 4; Length 938;
Best Local Similarity 81.8%; Pred. No. 3.7e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

Qy 1 ACCCATGATGGCGCATCTGATG 22
||| ||| ||| ||| ||| ||| |||
Db 76 ACCACTGATGGCTCAACTGATG 97

RESULT 30
US-09-883-985-1
; Sequence 1, Application US/09883985
; Patent No. 6635252
; GENERAL INFORMATION:
; APPLICANT: YU, GUO-LIANG
; ROSEN, CRAIG A.
; FRASER, CLAIRE M.
; GOCAYNE, JEANNINE D.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE-4
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, N.W., SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/883,985
; FILING DATE: 20-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/203,607
; FILING DATE: 02-DEC-1998
; APPLICATION NUMBER: US 08/722,050
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 08/225,757
; FILING DATE: 11-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1020003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (CDNA)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..879
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-883-985-1
Query Match 70.9%; Score 15.6; DB 4; Length 1080;
Best Local Similarity 81.8%; Pred. No. 3.8e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACCCATGATGGCGCATCTCATG 22
Db 502 AACCTGATGGAGCATCTCATG 523
RESULT 31
US-09-540-236-631/c
; Sequence 631, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 631
; LENGTH: 1236
; TYPE: DNA
; ORGANISM: M.cattarrhalis
US-09-540-236-631
Query Match 70.9%; Score 15.6; DB 4; Length 1236;
Best Local Similarity 81.8%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACCCATGATGGCGCATCTCATG 22
Db 571 ACCCATACGGCGGATTGATG 550
RESULT 32
US-09-270-767-11046/c
; Sequence 11046, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11046
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11046
Query Match 70.9%; Score 15.6; DB 4; Length 1292;
Best Local Similarity 81.8%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ACCCATGATGGCGCATCTCATG 22
Db 836 ATCGATGATGGCGCATCTCATG 815
RESULT 33
US-09-461-325-75/c
; Sequence 75, Application US/09461325A
; Patent No. 6475753
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/09/461,325A
; CURRENT FILING DATE: 1999-12-14
; EARLIER APPLICATION NUMBER: PCT/US99/13418
; EARLIER FILING DATE: 1999-06-15
; EARLIER APPLICATION NUMBER: 60/089,507
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,508
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,509
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/089,510
; EARLIER FILING DATE: 1998-06-16
; EARLIER APPLICATION NUMBER: 60/090,112
; EARLIER FILING DATE: 1998-06-22
; EARLIER APPLICATION NUMBER: 60/090,113
; EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 1650

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-325-75

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 4; Length 1650;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db 697 ACCCAGATGGCGCCTATGATG 676

RESULT 34
US-10-012-542-75/c
; Sequence 75, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-542-75

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 4; Length 1650;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db 697 ACCCAGATGGCGCCTATGATG 676

RESULT 35
US-10-115-123-75/c
; Sequence 75, Application US/10115123
; Patent No. 6774216
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029G30AP1D2
; CURRENT APPLICATION NUMBER: US/10/115,123
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: PCT/US99/13418
; PRIOR FILING DATE: 1999-06-15
; PRIOR APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089,510
```

```
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-115-123-75

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 4; Length 1650;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db 697 ACCCAGATGGCGCCTATGATG 676

RESULT 36
US-09-949-016-3778/c
; Sequence 3778, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3778
; LENGTH: 1939
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3778

Query Match
Best Local Similarity 70.9%; Score 15.6; DB 4; Length 1939;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db 1089 ACCCAGATGGCGCCTATGATG 1068

RESULT 37
US-09-540-236-425/c
; Sequence 425, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA-
; RALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 425
; LENGTH: 1962
; TYPE: DNA
; ORGANISM: M.catarrhalis
US-09-540-236-425
```

```
Query Match          70.9%; Score 15.6; DB 4; Length 1962;
Best Local Similarity 81.8%; Pred. No. 4.2e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
Db 1283 ACCCATGATGGCTCATCAAGTG 1262

RESULT 38
US-09-495-050A-214
; Sequence 214, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 214
; LENGTH: 2612
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2470285CT1
US-09-495-050A-214

Query Match          70.9%; Score 15.6; DB 4; Length 2612;
Best Local Similarity 81.8%; Pred. No. 4.4e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
Db 458 ATCCCTGATGGAGCATTTGATG 479

RESULT 39
US-08-409-995-3
; Sequence 3, Application US/08409995
; Patent No. 5646259
; GENERAL INFORMATION:
; APPLICANT: Barenkamp, Stephen I.
; APPLICANT: St. Geme III, Joseph W.
; TITLE OF INVENTION: Haemophilus Adhesion Proteins
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/409,995
; FILING DATE: 24-MAR-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-61053/RFT
```

```
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5738 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
US-08-409-995-3

Query Match          70.9%; Score 15.6; DB 1; Length 5738;
Best Local Similarity 81.8%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
Db 4174 ACAGATGATGGCGCAATGATG 4195

RESULT 40
US-08-685-467-3
; Sequence 3, Application US/08685467
; Patent No. 6060059
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Barenkamp, Stephen J.
; TITLE OF INVENTION: HAEMOPHILUS ADHESION PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,467
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/409,995
; FILING DATE: 24-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-61053-2/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5738 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-08-685-467-3

Query Match          70.9%; Score 15.6; DB 3; Length 5738;
Best Local Similarity 81.8%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGATG 22
Db 4174 ACAGATGATGGCGCAATGATG 4195
```

Search completed: June 4, 2005, 11:53:11
Job time : 50.9467 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 07:08:50 ; Search time 209.716 Seconds
(without alignments)
644.888 Million cell updates/sec

Title: US-09-674-277-24

Perfect score: 22

Sequence: 1 acccatggtggcgtctgatg 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
 - 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
 - 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.2	78.2	261	18	US-10-425-115-68733
2	17.2	78.2	479	18	US-10-425-115-81577
3	17.2	78.2	556	18	US-10-437-963-62826
4	17.2	78.2	975	17	US-10-166-653-5
5	17.2	78.2	1823	18	US-10-751-928-23
6	17.2	78.2	37051	16	US-10-004-113-55
7	16.8	76.4	659	17	US-10-260-238-1155
8	16.8	76.4	1003	18	US-10-425-115-41483
9	16.8	76.4	1078	17	US-10-424-599-90825
10	16.8	76.4	1560	18	US-10-437-963-60277
11	16.8	76.4	1629	17	US-10-369-493-41623

12	16.4	74.5	316	9	US-09-964-824A-342	Sequence 342, App
13	16.4	74.5	316	9	US-09-880-107-676	Sequence 676, App
14	16.4	74.5	316	19	US-10-843-641A-5645	Sequence 5645, Ap
15	16.4	74.5	329	18	US-10-425-115-164714	Sequence 164714,
16	16.4	74.5	352	18	US-10-425-115-151045	Sequence 151045,
c 17	16.4	74.5	600	9	US-09-823-245A-130	Sequence 130, App
18	16.4	74.5	654	18	US-10-767-701-10884	Sequence 10884, A
19	16.4	74.5	792	18	US-10-425-115-184238	Sequence 184238,
20	16.4	74.5	852	9	US-09-974-300-1199	Sequence 1199, Ap
21	16.4	74.5	1049	17	US-10-425-114-3453	Sequence 3453, Ap
22	16.4	74.5	1152	18	US-10-437-963-12578	Sequence 12578, A
23	16.4	74.5	1159	17	US-10-425-114-26319	Sequence 26319, A
24	16.4	74.5	1212	18	US-10-437-963-12580	Sequence 12580, A
c 25	16.4	74.5	1266	9	US-09-925-302-129	Sequence 129, App
c 26	16.4	74.5	1266	10	US-09-925-302-129	Sequence 129, App
27	16.4	74.5	1310	18	US-10-425-115-184237	Sequence 184237,
c 28	16.4	74.5	2065	14	US-10-103-313-60	Sequence 60, Appl
29	16.4	74.5	2121	18	US-10-425-115-184240	Sequence 184240,
c 30	16.4	74.5	2389	18	US-10-437-963-12577	Sequence 12577, A
c 31	16.4	74.5	2552	17	US-10-108-260A-390	Sequence 390, App
c 32	16.4	74.5	2766	18	US-10-437-963-12573	Sequence 12573, A
c 33	16.4	74.5	3240	18	US-10-437-963-78562	Sequence 78562, A
c 34	16.4	74.5	3621	18	US-10-437-963-12579	Sequence 12579, A
c 35	16.4	74.5	216929	18	US-10-741-601-5727	Sequence 5727, Ap
c 36	16.2	73.6	273	18	US-10-425-115-175495	Sequence 175495,
37	16.2	73.6	293	13	US-10-027-632-281143	Sequence 281143,
38	16.2	73.6	293	13	US-10-027-632-281144	Sequence 281144,
39	16.2	73.6	293	17	US-10-027-632-281143	Sequence 281143,
40	16.2	73.6	293	17	US-10-027-632-281144	Sequence 281144,
c 41	16.2	73.6	333	9	US-09-960-352-4722	Sequence 4722, Ap
c 42	16.2	73.6	366	18	US-10-425-115-182051	Sequence 182051,
c 43	16.2	73.6	394	18	US-10-425-115-29832	Sequence 29832, A
44	16.2	73.6	483	17	US-10-424-599-110586	Sequence 110586,
c 45	16.2	73.6	609	9	US-09-738-626-3337	Sequence 3337, Ap

ALIGNMENTS

RESULT 1
US-10-425-115-68733
; Sequence 68733, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 68733
; LENGTH: 261
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MET4577_162679C.1
US-10-425-115-68733

Query Match 78.2%; Score 17.2; DB 18; Length 261;
Best Local Similarity 86.4%; Pred. No. 1.8e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACCCATGATGGCGCATCTGATG 22
DB 72 ACCCATGATGGCGCATCTGATG 93
RESULT 2
US-10-425-115-81577

; Sequence 81577, Application US/10425115
; Publication No. US20040124272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 81577
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_174414C.1
US-10-425-115-81577

Query Match 78.2%; Score 17.2; DB 18; Length 479;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGATG 22
||| ||||| ||||| |||||
Db 207 ACAGATGATGGCACATCTGATG 228

RESULT 3
US-10-437-963-62826
; Sequence 62826, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 62826
; LENGTH: 556
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(556)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64127C.1
US-10-437-963-62826

Query Match 78.2%; Score 17.2; DB 18; Length 556;
Best Local Similarity 86.4%; Pred. No. 1.9e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGATG 22
||| ||||| ||||| |||||
Db 254 ACCAATGATGGGCACCTGTG 275

RESULT 4
US-10-166-653-5
; Sequence 5, Application US/10166653

; Publication No. US20030232338A1
; GENERAL INFORMATION:
; APPLICANT: USUDA, Yoshihiro
; APPLICANT: NISHIO, Yosuke
; APPLICANT: YASUEDA, Hisashi
; APPLICANT: SUGIMOTO, Shinichi
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING POLYPEPTIDES INVOLVED IN ONE-CARBON COM
; FILE REFERENCE: 211826USO
; CURRENT APPLICATION NUMBER: US/10/166,653
; CURRENT FILING DATE: 2002-06-12
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Methylophilus methylotrophus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(975)
; OTHER INFORMATION:
US-10-166-653-5

Query Match 78.2%; Score 17.2; DB 17; Length 975;
Best Local Similarity 86.4%; Pred. No. 2e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGATG 22
||| ||||| ||||| |||||
Db 54 ACCACTGGTGGGCATCTGATG 75

RESULT 5
US-10-751-928-23
; Sequence 23, Application US/10751928
; Publication No. US20040171134A1
; GENERAL INFORMATION:
; APPLICANT: Ajinomoto Co., Inc.
; TITLE OF INVENTION: Method for producing a recombinant strain of
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/751,928
; CURRENT FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: JP 2003-
; PRIOR FILING DATE: 2003-01-08
; FILE REFERENCE: OPI630
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 23
; LENGTH: 1823
; TYPE: DNA
; ORGANISM: Methylophilus methylotrophus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (522)..(1496)
US-10-751-928-23

Query Match 78.2%; Score 17.2; DB 18; Length 1823;
Best Local Similarity 86.4%; Pred. No. 2.1e+02;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGATG 22
||| ||||| ||||| |||||
Db 575 ACCACTGGTGGGCATCTGATG 596

RESULT 6
US-10-004-113-55
; Sequence 55, Application US/10004113
; Publication No. US20030194702A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; APPLICANT: Morris, David


```
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 60277
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_61820C.1
US-10-437-963-60277

Query Match          76.4%; Score 16.8; DB 18; Length 1560;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGA 20
Db 863 ACCGATGATGGCGCAGCTGA 882

RESULT 11
US-10-369-493-41623
; Sequence 41623, Application US/10369493
; Publication No. US2003033675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 41623
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: SPHINGOMONAS
US-10-369-493-41623

Query Match          76.4%; Score 16.8; DB 17; Length 1629;
Best Local Similarity 90.0%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTGAT 21
Db 631 CCGATGATGGCGCATCTGAT 650

RESULT 12
US-09-964-824A-342
; Sequence 342, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
; APPLICANT: Horrigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
```

```
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 342
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-342

Query Match          74.5%; Score 16.4; DB 9; Length 316;
Best Local Similarity 94.4%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTG 19
Db 178 CCCATGATGGCACATCTG 195

RESULT 13
US-09-880-107-676
; Sequence 676, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 676
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA284720
US-09-880-107-676

Query Match          74.5%; Score 16.4; DB 9; Length 316;
Best Local Similarity 94.4%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTG 19
Db 178 CCCATGATGGCACATCTG 195

RESULT 14
US-10-843-641A-5645
; Sequence 5645, Application US/10843641A
; Publication No. US20050064454A1
; GENERAL INFORMATION:
; APPLICANT: Avalon Pharmaceuticals, Inc.
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; TITLE OF INVENTION: Signature Gene Sets
; FILE REFERENCE: 689290-189
; CURRENT APPLICATION NUMBER: US/10/843,641A
```

; CURRENT FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/09/873,367
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US/09/954,531
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/962,832
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/09/964,824
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/09/967,768
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US/09/968,007
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,708
; PRIOR FILING DATE: 2001-10-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 8447
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5645
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-843-641A-5645

Query Match 74.5%; Score 16.4; DB 19; Length 316;
Best Local Similarity 94.4%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTG 19
DB 178 CCCATGATGGCGCATCTG 195

RESULT 15
US-10-425-115-164714
; Sequence 164714, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 164714
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_81799C.1
US-10-425-115-164714

Query Match 74.5%; Score 16.4; DB 18; Length 329;
Best Local Similarity 94.4%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ATGATGGCGCATCTGATG 22
DB 298 ATGATGGCGCATCTGATG 315

RESULT 16
US-10-425-115-151045

; Sequence 151045, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 151045
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_69271C.1
US-10-425-115-151045

Query Match 74.5%; Score 16.4; DB 18; Length 352;
Best Local Similarity 94.4%; Pred. No. 4.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 ATGATGGCGCATCTGATG 22
DB 124 ATGATGGCGCATCTGATG 141

RESULT 17
US-09-823-245A-130/c
; Sequence 130, Application US/09823245A
; Publication No. US20020039760A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtner, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalak
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6401
; CURRENT APPLICATION NUMBER: US/09/823,245A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/194,941
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 130
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-823-245A-130

Query Match 74.5%; Score 16.4; DB 9; Length 600;
Best Local Similarity 94.4%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTG 19
DB 456 CCCATGATGGCGCATCTG 439

RESULT 18
US-10-767-701-10884
; Sequence 10884, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 10884
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS75387_1
US-10-767-701-10884
```

```
Query Match          74.5%; Score 16.4; DB 18; Length 654;
Best Local Similarity 94.4%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 ATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db      259 ATGATGGCACATCTGATG 276
```

```
RESULT 19
US-10-425-115-184238
; Sequence 184238, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 184238
; LENGTH: 792
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_99612C.1
US-10-425-115-184238
```

```
Query Match          74.5%; Score 16.4; DB 18; Length 792;
Best Local Similarity 94.4%; Pred. No. 4.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 ATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db      769 ATGATGGCACATCTGATG 786
```

```
RESULT 20
US-09-974-300-1199
; Sequence 1199, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
```

```
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1199
; LENGTH: 852
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(852)
; OTHER INFORMATION: n = A,T,C or G
US-09-974-300-1199
```

```
Query Match          74.5%; Score 16.4; DB 9; Length 852;
Best Local Similarity 94.4%; Pred. No. 4.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 ATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db      363 ATGATGGCGCATCTGATG 380
```

```
RESULT 21
US-10-425-114-3453
; Sequence 3453, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 3453
; LENGTH: 1049
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700265457_FLI
US-10-425-114-3453
```

```
Query Match          74.5%; Score 16.4; DB 17; Length 1049;
Best Local Similarity 94.4%; Pred. No. 4.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      5 ATGATGGCGCATCTGATG 22
||||| ||||| ||||| |||||
Db      605 ATGATGGCACATCTGATG 622
```

```
RESULT 22
US-10-437-963-12578
; Sequence 12578, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
```

```
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 12578
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18693C.1
US-10-437-963-12578

Query Match          74.5%; Score 16.4; DB 18; Length 1152;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 CATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
Db      978 CATGATGGCGCATCTGAT 995

RESULT 23
US-10-425-114-26319
; Sequence 26319, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5321)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26319
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4148-010-D2_FLI
US-10-425-114-26319

Query Match          74.5%; Score 16.4; DB 17; Length 1159;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 ATGATGGCGCATCTGATG 22
      ||||| ||||| ||||| |||||
Db      762 ATGATGGCGCATCTGATG 779

RESULT 24
US-10-437-963-12580
; Sequence 12580, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 12580
; LENGTH: 1212
```

```
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_18693C.1
US-10-437-963-12580

Query Match          74.5%; Score 16.4; DB 18; Length 1212;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 CATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
Db      972 CATGATGGCGCATCTGAT 989

RESULT 25
US-09-925-302-129/c
; Sequence 129, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1222)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1235)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1243)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-129

Query Match          74.5%; Score 16.4; DB 9; Length 1266;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CCCATGATGGCGCATCTG 19
      ||||| ||||| ||||| |||||
Db      961 CCCATGATGGCGCATCTG 944

RESULT 26
US-09-925-302-129/c
; Sequence 129, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
```

```
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 129
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1222)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1235)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1243)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-129

Query Match          74.5%; Score 16.4; DB 10; Length 1266;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CCCATGATGGCGCATCTG 19
      |||||
Db      961 CCCATGATGGCGCATCTG 944

RESULT 27
US-10-425-115-184237
; Sequence 184237, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 184237
; LENGTH: 1310
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_99611C.1
US-10-425-115-184237

Query Match          74.5%; Score 16.4; DB 18; Length 1310;
Best Local Similarity 94.4%; Pred. No. 5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 ATGATGGCGCATCTGATG 22
      |||||
Db      796 ATGATGGCGCATCTGATG 813

RESULT 28
US-10-103-313-60/c
; Sequence 60, Application US/10103313
; Publication No. US20030082758A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653

; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 2065
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-103-313-60

Query Match          74.5%; Score 16.4; DB 14; Length 2065;
Best Local Similarity 94.4%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 CCCATGATGGCGCATCTG 19
      |||||
Db      1900 CCCATGATGGCGCATCTG 1883

RESULT 29
US-10-425-115-184240
; Sequence 184240, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 184240
; LENGTH: 2121
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_99614C.1
US-10-425-115-184240

Query Match          74.5%; Score 16.4; DB 18; Length 2121;
Best Local Similarity 94.4%; Pred. No. 5.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      5 ATGATGGCGCATCTGATG 22
      |||||
Db      1369 ATGATGGCGCATCTGATG 1386

RESULT 30
US-10-437-963-12577
; Sequence 12577, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 12577
; LENGTH: 2389
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
```

OTHER INFORMATION: Clone ID: PAT_MRT4530_18692C.1
US-10-437-963-12577

Query Match 74.5%; Score 16.4; DB 18; Length 2389;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGATGGCGCATCTGAT 21
|||||
DB 486 CATGATGGAGCATCTGAT 503

RESULT 31

US-10-108-260A-390/c
Sequence 390, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1el full length cDNA
FILE REFERENCE: H1-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 390
LENGTH: 2552
TYPE: DNA
ORGANISM: Homo sapiens
US-10-108-260A-390

Query Match 74.5%; Score 16.4; DB 17; Length 2552;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTG 19
|||||
DB 2410 CCCATGATGGAGCATCTG 2393

RESULT 32

US-10-437-963-12573
Sequence 12573, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 12573
LENGTH: 2766
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_18689C.1
US-10-437-963-12573

Query Match 74.5%; Score 16.4; DB 18; Length 2766;
Best Local Similarity 94.4%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGATGGCGCATCTGAT 21
|||||
DB 1614 CATGATGGAGCATCTGAT 1631

RESULT 33

US-10-437-963-78562/c
Sequence 78562, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 78562
LENGTH: 3240
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_78357C.1
US-10-437-963-78562

Query Match 74.5%; Score 16.4; DB 18; Length 3240;
Best Local Similarity 94.4%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGATGGCGCATCTGAT 21
|||||
DB 956 CATGATGGAGCATCTGAT 939

RESULT 34

US-10-437-963-12579
Sequence 12579, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 12579
LENGTH: 3621
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_18694C.1
US-10-437-963-12579

Query Match 74.5%; Score 16.4; DB 18; Length 3621;
Best Local Similarity 94.4%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CATGATGGCGCATCTGAT 21
|||||
DB 2583 CATGATGGAGCATCTGAT 2600

```
RESULT 35
US-10-741-601-5727/c
; Sequence 5727, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5727
; LENGTH: 216929
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(216929)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-601-5727

Query Match 74.5%; Score 16.4; DB 18; Length 216929;
Best Local Similarity 94.4%; Pred. No. 7.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 CCATGATGGCGCATCTGTA 20
||||| |||||||
Db 47746 CCATGAAGCGCATCTGTA 47729

RESULT 36
US-10-425-115-175495/c
; Sequence 175495, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 175495
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MPT4577_91639C.1
US-10-425-115-175495

Query Match 73.6%; Score 16.2; DB 18; Length 273;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| |||||||
Db 135 ACCCAAGATGGTGCATCTGTT 115

RESULT 37
US-10-027-632-281143
; Sequence 281143, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281143
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281143

Query Match 73.6%; Score 16.2; DB 13; Length 293;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| |||||||
Db 224 AACCAAGATGGTGCATCTGAT 244
```

```
FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281143
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281143

Query Match 73.6%; Score 16.2; DB 13; Length 293;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| |||||||
Db 224 AACCAAGATGGTGCATCTGAT 244
```

```
RESULT 38
US-10-027-632-281144
; Sequence 281144, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281144
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281144

Query Match 73.6%; Score 16.2; DB 13; Length 293;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| |||||||
Db 224 AACCAAGATGGTGCATCTGAT 244
```


Db 224 AACCAAGATGGTGCATCTGAT 244

RESULT 39

US-10-027-632-281143
; Sequence 281143, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281143
; LENGTH: 293
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281143

Query Match 73.6%; Score 16.2; DB 17; Length 293;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
Db 224 AACCAAGATGGTGCATCTGAT 244

RESULT 40

US-10-027-632-281144
; Sequence 281144, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 281144

; LENGTH: 293
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-281144

Query Match 73.6%; Score 16.2; DB 17; Length 293;
Best Local Similarity 85.7%; Pred. No. 5.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
Db 224 AACCAAGATGGTGCATCTGAT 244

Search completed: June 4, 2005, 12:20:00
Job time : 210.716 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 4, 2005, 06:32:00 ; Search time 1343.17 Seconds
(without alignments)
623.460 Million cell updates/sec

Title: US-09-674-277-24

Perfect score: 22

Sequence: 1 acccatgatggcgcattgatg 22

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hcc:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gse1:*
- 9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19.4	88.2	907	CG021934	ZMMBBc055
2	18.8	85.5	542	BE334902	us90601.y
3	18.8	85.5	585	AW592571	hf44609.x
4	17.8	80.9	244	CL220229	ZMMBBc046
5	17.8	80.9	258	CG020344	ZMMBBc055
6	17.8	80.9	258	CG914782	ZMMBBb037
7	17.8	80.9	297	BZ720255	PUCV23TD
8	17.8	80.9	300	CC159184	if99a04.b
9	17.8	80.9	319	CC600945	ZMMBBc040
10	17.8	80.9	326	CE207422	tigr-gss-
11	17.8	80.9	398	CC176034	ZMMBBc029
12	17.8	80.9	410	CG900788	ZMMBBb051
13	17.8	80.9	469	CG913499	ZMMBBb037
14	17.8	80.9	484	CC891160	ZMMBBb051
15	17.8	80.9	507	CR587491	CR587491
16	17.8	80.9	524	BH775554	fzmb011fo
17	17.8	80.9	544	CC876552	ZMMBBb019
18	17.8	80.9	557	CM005355	ZMMBla000
19	17.8	80.9	562	CC792074	ZMMBBb016
20	17.8	80.9	574	CG396746	OG4BF957C
21	17.8	80.9	576	BZ535002	OGAEJ367C
22	17.8	80.9	601	CG970056	ZUAD233TV
23	17.8	80.9	609	CA241511	SCUTFL307
24	17.8	80.9	614	CG289521	OGZAD26TH

25	17.8	80.9	631	9	CG705950
c 26	17.8	80.9	656	9	CG851685
c 27	17.8	80.9	677	9	CG880804
28	17.8	80.9	684	8	CC460558
c 29	17.8	80.9	689	9	CC742896
30	17.8	80.9	690	9	CC803004
c 31	17.8	80.9	691	8	BZ998359
c 32	17.8	80.9	693	9	CC629672
33	17.8	80.9	697	8	CC464996
c 34	17.8	80.9	697	9	CG913717
c 35	17.8	80.9	698	8	CC159056
c 36	17.8	80.9	703	9	CC744522
c 37	17.8	80.9	704	9	CC788457
c 38	17.8	80.9	708	9	CG352166
c 39	17.8	80.9	711	9	CC876585
c 40	17.8	80.9	717	9	CC750113
c 41	17.8	80.9	717	9	CG793584
c 42	17.8	80.9	730	9	CG915625
c 43	17.8	80.9	737	9	CG683300
c 44	17.8	80.9	749	9	CC894497
45	17.8	80.9	751	8	BZ534586

ALIGNMENTS

RESULT 1
CG021934
LOCUS
DEFINITION
ZMMBBc0555p17r ZMMBBc Zea mays genomic clone ZMMBBc0555p17 3',
genomic survey sequence.
ACCESSION
CG021934
VERSION
CG021934.1
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Zea mays
REFERENCE
1 (bases 1 to 907)
AUTHORS
Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Rouzard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE
Sequencing of the maize genome at PGIR (2003b)
JOURNAL
Unpublished (2003)
COMMENT
Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Wakeman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@wakeman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 132.
Location/Qualifiers
1..907
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBc0555p17"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBBc"
/note="Vector: pTABAC1.3; Site_1: BamHI; Site_2: BamHI"

Query Match 88.2%; Score 19.4; DB 9; Length 907;
Best Local Similarity 95.2%; Pred. No. 1.4e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACCCATGATGGGCATCTGAT 21
|||||

```

Db      756 ACCCATGATGGCGCATCTCGAT 776

RESULT 2
BE334902
LOCUS   BE334902          542 bp    mRNA    linear    EST 14-JUL-2000
DEFINITION us90e01.y1 Soares mammary gland NMLMG Mus musculus cDNA clone
IMAGE:3325656 5', similar to TR:O14618 O14618 COPPER CHAPERONE FOR
SUPEROXIDE DISMUTASE. ;, mRNA sequence.

ACCESSION BE334902
VERSION   BE334902.1 GI:9208678
KEYWORDS  Mus musculus (house mouse)
SOURCE    Mus musculus
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 542)
AUTHORS   Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE     Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Mus.
JOURNAL   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT   National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco
MG1:1069820
High quality sequence stop: 461.
Location/Qualifiers
1..542
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3325656"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NMLMG"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match      85.5%; Score 18.8; DB 2; Length 542;
Best Local Similarity 90.9%; Pred. No. 2.6e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ACCCATGATGGCGCATCTCGATG 22
|||||
Db      516 ACCCTGATGGAGCATCTCGATG 537
|||||

RESULT 3
AW592571/c
LOCUS   AW592571          585 bp    mRNA    linear    EST 22-MAR-2000
DEFINITION hf44e09.x1 Soares NFL T GBC S1 Homo sapiens cDNA clone
IMAGE:2934760 3', similar to TR:O14618 O14618 COPPER CHAPERONE FOR
SUPEROXIDE DISMUTASE. ;, mRNA sequence.

ACCESSION AW592571
VERSION   AW592571.1 GI:7279755
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 585)
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LML; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40RP from Gibco
MG1:1069820
High quality sequence stop: 461.
Location/Qualifiers
1..542
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:3325656"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/clone_lib="Soares mammary gland NMLMG"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7T3 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match      85.5%; Score 18.8; DB 2; Length 542;
Best Local Similarity 90.9%; Pred. No. 2.6e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 ACCCATGATGGCGCATCTCGATG 22
|||||
Db      516 ACCCTGATGGAGCATCTCGATG 537
|||||

RESULT 4
LOCUS   CL220229/c
DEFINITION ZMMBBc0460A05f ZMMBBc Zea mays genomic clone ZMMBBc0460A05 5',
genomic survey sequence.
ACCESSION CL220229
VERSION   CL220229.1 GI:40752903
KEYWORDS  Zea mays
SOURCE    Zea mays
ORGANISM  Zea mays
REFERENCE 1 (bases 1 to 244)
AUTHORS   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
TITLE     Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
JOURNAL   Sequencing of the maize genome at PGIR (2003c)
COMMENT   Unpublished (2003)
Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: 17
Class: BAC ends.
Location/Qualifiers
1..244
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"

FEATURES
source

```

```

/clone="ZMMBBc0460A05"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBBc"
/note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 244;
Best Local Similarity 90.5%; Pred. No. 7.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 114 ACCCAAGATGGTGCATCTGAT 94

RESULT 5
CG020344
LOCUS ZMMBBc0551E16r ZMMBBc Zea mays genomic clone ZMMBBc0551E16 3',
DEFINITION genomic survey sequence.
ACCESSION CG020344
VERSION CG020344.1 GI:33892509
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 258)
AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Rouard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE Sequencing of the maize genome at PGR (2003b)
JOURNAL Unpublished (2003)
COMMENT Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 79.
Location/Qualifiers
1..258
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBc0551E16"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBBc"
/note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 258;
Best Local Similarity 90.5%; Pred. No. 7.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 72 ACCCAAGATGGTGCATCTGAT 92

RESULT 6
CG914782/c
LOCUS ZMMBBb0377M18.r ZMMBBb Zea mays genomic clone ZMMBBb0377M18 3',
DEFINITION genomic survey sequence.
ACCESSION CG914782
VERSION CG914782.1 GI:39614068
KEYWORDS GSS.

```

```

SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 258)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0377 row: M column: 18
Seq primer: M13r
Class: BAC ends.
Location/Qualifiers
1..258
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBb0377M18"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII; Zea mays L. ssp. mays"

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 258;
Best Local Similarity 90.5%; Pred. No. 7.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 174 ACCCAAGATGGTGCATCTGAT 154

RESULT 7
BZ720255
LOCUS PUCFV23TD ZM 0.6_1.0_KB Zea mays genomic clone ZMMBTal30C22,
DEFINITION genomic survey sequence.
ACCESSION BZ720255
VERSION BZ720255.1 GI:28510221
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 297)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.

```

Match

JOURNAL Science 301 (5641), 1898-1903 (2003)
 MEDLINE 22875432
 PUBMED 14512627
 COMMENT Contact: Kirkness EF
 The Institute for Genomic Research
 Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
 Rockville, MD 20850, USA
 Tel: 301-838-0200
 Fax: 301-838-0208
 Email: ekirknes@tigr.org
 Class: shotgun.

Location/Qualifiers
 1..326
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /note="Site 1: BetXI; Libraries were prepared from
 peripheral blood"

FEATURES

source

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 326;
 Best Local Similarity 90.5%; Pred. No. 7.5e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTGAT 22

|||||

Db 166 CCCATGATGCCACATCTGAT 146

RESULT 11

CC176034

LOCUS

DEFINITION CC176034 398 bp DNA linear GSS 02-MAY-2003
 ZMMBBc0297L18r ZMMBBc Zea mays genomic clone ZMMBBc0297L18 3',
 genomic survey sequence.

ACCESSION CC176034

VERSION CC176034.1

KEYWORDS GI:30275806

SOURCE GSS.

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 398)

Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
 Sugiyama, R., Fuks, G., Yu, Y., Wing, R. and Messing, J.

Sequencing of the maize genome at PGIR (2003a)

Unpublished (2003)

Contact: Bharti, A.K.

Dr. Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Wakeman Institute, Rutgers

University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@wakeman.rutgers.edu

Seq primer: SP6

Class: BAC ends

High quality sequence start: 93.

Location/Qualifiers

1..398

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZMMBBc0297L18"

/lab_host="E. coli DH10B"

/clone_lib="ZMMBBc"

/note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN

Query Match 80.9%; Score 17.8; DB 8; Length 398;

Best Local Similarity 90.5%; Pred. No. 7.7e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21

|||||

Db 185 ACCCAAGATGGTGCATCTGAT 205

RESULT 12

CG900788/c

LOCUS

DEFINITION CG900788 410 bp DNA linear GSS 09-DEC-2003
 ZMMBBb0511J08f ZMMBBb (HindIII) Zea mays genomic clone
 ZMMBBb0511J08 5', genomic survey sequence.

ACCESSION CG900788

VERSION CG900788.1

KEYWORDS GI:39600055

SOURCE GSS.

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 410)

Bharti, A.K., Young, S., Kavchok, S., Keizer, G., Bronzino, A.C.,
 Zohovetz, V., Fuks, G., Yu, Y., Wing, R. and Messing, J.

Sequencing of the maize genome at PGIR (2003c)

Unpublished (2003)

Contact: Bharti, A.K.

Dr. Joachim Messing's lab

The Plant Genome Initiative at Rutgers, Wakeman Institute, Rutgers

University

190 Frelinghuysen Road, Piscataway, NJ 08854, USA

Tel: 732 445 3801

Fax: 732 445 5735

Email: bharti@wakeman.rutgers.edu

Seq primer: T7

Class: BAC ends

High quality sequence start: 138.

Location/Qualifiers

1..410

/organism="Zea mays"

/mol_type="genomic DNA"

/cultivar="B73"

/db_xref="taxon:4577"

/clone="ZMMBBb0511J08"

/lab_host="E. coli DH10B"

/clone_lib="ZMMBBb (HindIII)"

/note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 410;

Best Local Similarity 90.5%; Pred. No. 7.7e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21

|||||

Db 377 ACCCAAGATGGTGCATCTGAT 357

RESULT 13

CG913499/c

LOCUS

DEFINITION CG913499 469 bp DNA linear GSS 09-DEC-2003
 ZMMBBb0373J04.r ZMMBBb Zea mays genomic clone ZMMBBb0373J04 3',
 genomic survey sequence.

ACCESSION CG913499

VERSION CG913499.1

KEYWORDS GI:39612785

SOURCE GSS.

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 469)

Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.

REFERENCE

TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu

PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0373 row: J column: 04
Seq primer: M13r
Class: BAC ends.

FEATURES

source
1..469
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBb0373J04"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. ssp. mays"

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 469;
Best Local Similarity 90.5%; Pred. No. 7.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTCGAT 21
||||| ||||| ||||| ||||| |||||
Db 441 ACCCAAGATGGTGCATCTGAT 421

RESULT 14
LOCUS CC8911160
DEFINITION ZMMBBc0516K23f ZMMBBc Zea mays genomic clone ZMMBBc0516K23 5',
genomic survey sequence.
ACCESSION CC8911160
VERSION CC8911160.1 GI:33370587
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 484)
Bharti.A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Rouard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
Sequencing of the maize genome at PGIR (2003b)
Unpublished (2003)
Contact: Bharti,A.K.

REFERENCE
AUTHORS Dr.Joachim Messing's lab
TITLE The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
JOURNAL University
COMMENT 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 123.

FEATURES

source
1..484
Location/Qualifiers
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"

/db_xref="taxon:4577"
/clone="ZMMBBc0516K23"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBBc"
/note="Vector: pTARBAC1.3; Site 1: BamHI; Site 2: BamHI"

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 484;
Best Local Similarity 90.5%; Pred. No. 7.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 209 ACCCAAGATGGTGCATCTGAT 229

RESULT 15

LOCUS CR587491
DEFINITION CR587491 XGC-tailbud-head Xenopus tropicalis cDNA clone THdA022j18
3', mRNA sequence.
ACCESSION CR587491
VERSION CR587491.1 GI:50587491
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.

REFERENCE
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
COMMENT Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: THdA022j18.q1kT7
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
Seq primer: T7.

FEATURES

source
1..507
Location/Qualifiers
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="THdA022j18"
/dev_stage="tailbud head (stage 28-30)"
/lab_host="Escherichia coli DH10B."
/clone_lib="XGC-tailbud-head"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from tailbud
head. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

Query Match 80.9%; Score 17.8; DB 7; Length 507;
Best Local Similarity 90.5%; Pred. No. 7.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CCCATGATGGCGCATCTGAT 22
||||| ||||| ||||| ||||| |||||
Db 453 CCCCTGATGGCACATCTGATG 473

RESULT 16

BH775554/c
LOCUS BH775554
DEFINITION fzm011f013h10f0 fzmdb filtered library Zea mays genomic clone
fzm011f013h10 5', genomic survey sequence.
ACCESSION BH775554
VERSION BH775554.1 GI:19777741


```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 524)
Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.
GeneThresher methylation filtered genomic sequences from maize
Unpublished (2002)
Contact: Bedell JA
Orion Genomics, LLC
4041 Forest Park Ave, St. Louis, MO 63108, USA
Tel: 314 615 6979
Fax: 314 615 5975
Email: jbedell@oriongenomics.com
Plate: f2mb011f013 row: h column: 10
Seq primer: M13 forward
Class: shotgun
High quality sequence stop: 524.
Location/Qualifiers
1..524
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="Mol1"
/db_xref="taxon:4577"
/clone="f2mb011f013h10"
/notes="Organ: leaf; Vector: pBCSK(-); Site 1: HincII; DNA
prepared from purified nuclei was randomly sheared,
end-repaired, size fractionated to enrich for the 0.5 to
5 kb fraction, ligated into HincII-digested pBCSK(-)
vector and electroporated into E. coli cells."

FEATURES
source
1 ACCCATGATGGCGCATCTGAT 21
|||||
43 ACCCAAGATGGTGCACTCTGAT 23

ORIGIN
Query Match 80.9%; Score 17.8; DB 8; Length 524;
Best Local Similarity 90.5%; Pred. No. 8e+02; 2; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
|||||
43 ACCCAAGATGGTGCACTCTGAT 23

Db 1 ACCCATGATGGCGCATCTGAT 21
|||||
43 ACCCAAGATGGTGCACTCTGAT 23

RESULT 17
CC876552/c
LOCUS
DEFINITION
ZMBBB0195P23.f ZMBBB Zea mays genomic clone ZMBBB0195P23 5',
genomic survey sequence.
ACCESSION
CC876552
VERSION
CC876552.1 GI:33306645
KEYWORDS
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 544)
Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
Sequencing of the maize genome
Unpublished (2003)
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: T7
BACKWARD: M13r

Plate: 0195 row: F column: 23
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..544
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMBBB0195P23"
/lab_host="DH10B"
/clone_lib="ZMBBB"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII; Zea mays L. ssp. mays"

FEATURES
source
1 ACCCATGATGGCGCATCTGAT 21
|||||
521 ACCCAAGATGGTGCACTCTGAT 501

ORIGIN
Query Match 80.9%; Score 17.8; DB 9; Length 544;
Best Local Similarity 90.5%; Pred. No. 8e+02; 2; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
|||||
521 ACCCAAGATGGTGCACTCTGAT 501

Db 1 ACCCATGATGGCGCATCTGAT 21
|||||
521 ACCCAAGATGGTGCACTCTGAT 501

RESULT 18
CW005355
LOCUS
DEFINITION
ZMBLA0005118.r ZMBLA Zea mays genomic clone ZMBLA0005118 3',
genomic survey sequence.
ACCESSION
CW005355
VERSION
CW005355.1 GI:52583930
KEYWORDS
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 557)
Wing,R., Luo,M., Soderlund,C. and Haller,K.
ZMML sequences
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
Plate: 0005 row: I column: 18
Class: BAC ends.
Location/Qualifiers
1..557
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMBLA0005118"
/tissue_type="immature ears"
/lab_host="DH10B T1 phage resistant"
/clone_lib="ZMBLA"
/notes="Vector: pAGIBAC1; Site_1: SalI; Site_2: SalI"

FEATURES
source
1 ACCCATGATGGCGCATCTGAT 21
|||||
226 ACCCAAGATGGTGCACTCTGAT 246

ORIGIN
Query Match 80.9%; Score 17.8; DB 9; Length 557;
Best Local Similarity 90.5%; Pred. No. 8e+02; 2; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
|||||
226 ACCCAAGATGGTGCACTCTGAT 246

Db 1 ACCCATGATGGCGCATCTGAT 21
|||||
226 ACCCAAGATGGTGCACTCTGAT 246

```

```
RESULT 19
CC792074      562 bp   DNA      linear      GSS 01-JUL-2003
LOCUS        ZM535002.1 ZM535002.1 genomic clone ZM535002.1
DEFINITION   genomic survey sequence.
ACCESSION    CC792074
VERSION      CC792074.1 GI:32387297
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 562)
AUTHORS      Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.
              and Wing, R.
TITLE        Sequencing of the maize genome
JOURNAL      Unpublished (2003)
COMMENT      Contact: Rod Wing
              Arizona Genomics Institute
              University of Arizona
              Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
              85721-0088, USA
              Tel: 520 626 3967
              Fax: 520 621 9288
              Email: http://genome.arizona.edu
              PCR Primers
              FORWARD: T7
              BACKWARD: M13r
              Plate: 0165 row: C column: 23
              Seq primer: M13r
              Class: BAC ends.
              Location/Qualifiers
                1..562
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /cultivar="B73"
                /db_xref="taxon:4577"
                /clone="ZM535002.1"
                /lab_host="DH10B"
                /clone_lib="ZM535002.1"
                /note="Vector: pBAC11; Site_1: HindIII; Site_2:
                HindIII; Zea mays L. ssp. mays"

FEATURES             source
  source
    Query Match      80.9%; Score 17.8; DB 9; Length 562;
    Best Local Similarity 90.5%; Pred. No. 8e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 316 ACCCAAGATGGTGCATCTGAT 296

RESULT 20
CG296746/7c     574 bp   DNA      linear      GSS 25-AUG-2003
LOCUS        CG296746/7c ZM 0.7 1.5 KB Zea mays genomic clone ZM535002.1
DEFINITION   genomic survey sequence.
ACCESSION    CG296746
VERSION      CG296746.1 GI:34210960
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 574)
AUTHORS      Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
              Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
              Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
              Consortium for Maize Genomics
              Unpublished (2002)
TITLE        Consortium for Maize Genomics
JOURNAL      Unpublished (2002)

COMMENT      Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: T7
              Class: sheared ends.
              Location/Qualifiers
                1..574
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZM535002.1"
                /clone_lib="ZM 0.7 1.5 KB"
                /note="Vector: pBAC11; Site_1: HindIII; Site_2:
                HindIII; Zea mays L. ssp. mays"

FEATURES             source
  source
    Query Match      80.9%; Score 17.8; DB 8; Length 576;
    Best Local Similarity 90.5%; Pred. No. 8.1e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 231 ACCCAAGATGGTGCATCTGAT 211

RESULT 21
BZ535002        576 bp   DNA      linear      GSS 16-DEC-2002
LOCUS        BZ535002 ZM2 0.7 1.5 KB Zea mays genomic clone ZM535002.1
DEFINITION   genomic survey sequence.
ACCESSION    BZ535002
VERSION      BZ535002.1 GI:27081645
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 576)
AUTHORS      Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
              Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
              Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
              Consortium for Maize Genomics
              Unpublished (2002)
TITLE        Consortium for Maize Genomics
JOURNAL      Unpublished (2002)
COMMENT      Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: T7
              Class: sheared ends.
              Location/Qualifiers
                1..576
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZM535002.1"
                /clone_lib="ZM2 0.7 1.5 KB"
                /note="Vector: pBAC11; Site_1: HindIII; Site_2:
                HindIII; Zea mays L. ssp. mays"

FEATURES             source
  source
    Query Match      80.9%; Score 17.8; DB 8; Length 576;
    Best Local Similarity 90.5%; Pred. No. 8.1e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 231 ACCCAAGATGGTGCATCTGAT 211

RESULT 22
CG296746/7c     574 bp   DNA      linear      GSS 25-AUG-2003
LOCUS        CG296746/7c ZM 0.7 1.5 KB Zea mays genomic clone ZM535002.1
DEFINITION   genomic survey sequence.
ACCESSION    CG296746
VERSION      CG296746.1 GI:34210960
KEYWORDS     GSS.
SOURCE       Zea mays
ORGANISM     Zea mays
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
              clade; Panicoideae; Andropogoneae; Zea.
REFERENCE    1 (bases 1 to 574)
AUTHORS      Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
              Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
              Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
              Consortium for Maize Genomics
              Unpublished (2002)
TITLE        Consortium for Maize Genomics
JOURNAL      Unpublished (2002)

COMMENT      Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: T7
              Class: sheared ends.
              Location/Qualifiers
                1..574
                /organism="Zea mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZM535002.1"
                /clone_lib="ZM 0.7 1.5 KB"
                /note="Vector: pBAC11; Site_1: HindIII; Site_2:
                HindIII; Zea mays L. ssp. mays"

FEATURES             source
  source
    Query Match      80.9%; Score 17.8; DB 8; Length 576;
    Best Local Similarity 90.5%; Pred. No. 8.1e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 231 ACCCAAGATGGTGCATCTGAT 211
```

Db 312 ACCCAAGTGGTGCACTCGAT 332
||||| ||||| ||||| ||||| |||||

RESULT 22
CC970056/c 601 bp DNA linear GSS 18-AUG-2003
LOCUS ZUAD233TV_ZM_3.0_4.0_KB Zea mays genomic clone ZMMPa0037F18,
DEFINITION genomic survey sequence.

ACCESSION CC970056
VERSION CC970056
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

AUTHORS 1 (bases 1 to 601)

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics
Unpublished (2002)

Other GSSs: ZUAD233TH
Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

FEATURES
source Location/Qualifiers
1..601

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone_lib="ZMMPa0037F18"

/clone_lib="ZM_3.0_4.0_KB"

/notes="Vector: pBCSK-; Site_1: HincII; 3-4 kb 'unfiltered'
genomic DNA library"

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 601;

Best Local Similarity 90.5%; Pred.No. 8.1e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGCGCATCTGAT 21

||||| ||||| ||||| ||||| |||||

Db 324 ACCCAAGTGGTGCACTCGAT 304

RESULT 23

CA241511/c

LOCUS SCUTFL3075H11.1 609 bp mRNA linear EST 25-SEP-2003
DEFINITION cDNA clone SCUTFL3075H11 5', mRNA sequence.

ACCESSION CA241511

VERSION CA241511.1 GI:35316195

KEYWORDS EST.

SOURCE Saccharum officinarum

ORGANISM Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum

complex.

REFERENCE 1 (bases 1 to 609)

Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.

The libraries that made SUCEST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bccccenter.fcav.unesp.br

Plate: 075 row: H column: 11

Seq primer: T7 Promoter Primer.

Location/Qualifiers

1..609

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCUTFL3075H11"

/lab_host="DH10B"

/clone_lib="Saccharum officinarum FL3"

/note="Organ: Base of developing inflorescence (5cm-long);

Vector: pSport1; Site_1: SalI; Site_2: NotI; An

unidirectional cDNA library generated from [Base of

developing inflorescence (5cm-long)]. cDNA was prepared

from polyA+ mRNA using SuperScript Plasmid System Kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 Kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 80.9%; Score 17.8; DB 6; Length 609;

Best Local Similarity 90.5%; Pred.No. 8.1e+02;

Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCCATGATGGCGCATCTGAT 21

||||| ||||| ||||| ||||| |||||

Db 76 ACCCATGAGCGCATCTGAT 56

RESULT 24

CG289521

LOCUS OGZAD26TH_ZM_0.7_1.5_KB Zea mays genomic clone ZMMPa0654F04,

DEFINITION genomic survey sequence.

ACCESSION CG289521

VERSION CG289521.1 GI:34203735

KEYWORDS GSS.

SOURCE Zea mays

ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

REFERENCE 1 (bases 1 to 614)

Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,

Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,

Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OGZAD26TV

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

FEATURES
source Location/Qualifiers
1..614

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

```

/db_xref="taxon:4577"
/clone="ZMBB0340H02"
/clone_lib="ZM 0.7 1.5 KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 614;
Best Local Similarity 90.5%; Pred. No. 8.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 549 ACCCAAGATGGTGCATCTGAT 569

RESULT 25
CG705950
LOCUS
DEFINITION
01S0621-04C1-C12 UniformMu MUTAIL Library Zea mays genomic clone
01S0621-04C1-C12, genomic survey sequence.
ACCESSION
CG705950
VERSION
CG705950.1 GI:37696560
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 631)
AUTHORS
Latshaw,S., Tan,B.-C., Settles,A.M. and McCarty,D.R.
TITLE
Sequence tagged transposon insertions from the UniformMu maize
population
JOURNAL
Unpublished (2003)
COMMENT
Contact: Donald R. McCarty
Plant Molecular and Cellular Biology Program
University of Florida
PO 110690 Gainesville, FL 32611-0690, USA
Tel: 352-392-1928 x322
Email: drmc@ufl.edu
Sequence flanking probable Mu insertion site in UniformMu line:
01S0621-04. Primer set: C
Class: transposon insertion site.
FEATURES
source
Location/Qualifiers
1..631
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="W22 (ACR, bz1-m9)"
/cultivar="UniformMu"
/db_xref="taxon:4577"
/clone="01S0621-04C1-C12"
/clone_lib="UniformMu MUTAIL Library"
/notes="Vector: TOPO-PCR4; DNA flanking Mu transposon
insertions in Mu inactive lines were extracted from the
UniformMu maize population by the thermo asymmetric
interlaced PCR (TAIL) protocol using primers specific for
the Mu terminal inverted repeat and a set of 16 arbitrary
primers. Amplicons were size enriched using Sepharose 400
spin columns and cloned into the TOPO PCR4 vector."

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 611;
Best Local Similarity 90.5%; Pred. No. 8.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 171 ACCCAAGATGGTGCATCTGAT 191

RESULT 26
CG851685/c
LOCUS

```

```

ZMBB0340H02.r ZMBB0340H02 Zea mays genomic clone ZMBB0340H02 3',
genomic survey sequence.
ACCESSION
CG851685
VERSION
CG851685.1 GI:38378546
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 656)
AUTHORS
Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE
Sequencing of the maize genome
JOURNAL
Unpublished (2003)
COMMENT
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0340 row: H column: 02
Seq primer: M13r
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..656
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMBB0340H02"
/lab_host="DH10B"
/clone_lib="ZMBB0340H02"
/notes="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. sep. mays"

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 656;
Best Local Similarity 90.5%; Pred. No. 8.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 401 ACCCAAGATGGTGCATCTGAT 381

RESULT 27
CG880804/c
LOCUS
DEFINITION
ZMBB0218M19.r ZMBB0340H02 Zea mays genomic clone ZMBB0218M19 3',
genomic survey sequence.
ACCESSION
CG880804
VERSION
CG880804.1 GI:33315490
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 677)
AUTHORS
Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE
Sequencing of the maize genome
JOURNAL
Unpublished (2003)
COMMENT
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

```

```

85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
  FORWARD: T7
  BACKWARD: M13r
Plate: 0218 row: M column: 19
Seq primer: M13r
Class: BAC ends.

FEATURES
  source
    1. .677
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /cultivar="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBBb0218M19"
    /lab_host="DH10B"
    /clone_lib="ZMMBBb"
    /note="Vector: pBeloBAC1.1; Site_1: HindIII; Site_2: HindIII; Zea mays L. ssp. mays"

ORIGIN
  Query Match      80.9%; Score 17.8; DB 9; Length 677;
  Best Local Similarity 90.5%; Pred. NO. 8.2e+02;
  Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 432 ACCAAGATGGTGCATCTGAT 412

RESULT 28
CC460558 684 bp DNA linear GSS 04-JUN-2003
LOCUS ZMMBBc0381K02r ZMMBBc Zea mays genomic clone ZMMBBc0381K02 3',
DEFINITION genomic survey sequence.
ACCESSION CC460558
VERSION CC460558.1 GI:31378147
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 684)
AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Rouzaud,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE Sequencing of the maize genome at PGIR (2003b)
JOURNAL Unpublished (2003)
COMMENT Contact: Bharti,A.K.
Dr. Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 398.
Location/Qualifiers
  1. .684
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /cultivar="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBBc0381K02"
  /lab_host="E. coli DH10B"
  /clone_lib="ZMMBBc"
  /note="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

ORIGIN
  Query Match      80.9%; Score 17.8; DB 9; Length 689;
  Best Local Similarity 90.5%; Pred. NO. 8.3e+02;
  Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 584 ACCAAGATGGTGCATCTGAT 564

RESULT 30
CC803004 690 bp DNA linear GSS 14-JUL-2003
LOCUS ih10b03.g1 WGS-ZmaysU (DH10b or JM107MA2) Zea mays genomic clone
DEFINITION ih10b03, genomic survey sequence.
ACCESSION CC803004
VERSION CC803004.1 GI:32627260
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 689)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
  FORWARD: T7
  BACKWARD: M13r
Plate: 0117 row: C column: 08
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
  1. .689
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /cultivar="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBBb0117C08"
  /lab_host="DH10B"
  /clone_lib="ZMMBBb"
  /note="Vector: pBeloBAC1.1; Site_1: HindIII; Site_2: HindIII; Zea mays L. ssp. mays"

ORIGIN
  Query Match      80.9%; Score 17.8; DB 9; Length 689;
  Best Local Similarity 90.5%; Pred. NO. 8.3e+02;
  Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 584 ACCAAGATGGTGCATCTGAT 564

RESULT 30
CC803004 690 bp DNA linear GSS 14-JUL-2003
LOCUS ih10b03.g1 WGS-ZmaysU (DH10b or JM107MA2) Zea mays genomic clone
DEFINITION ih10b03, genomic survey sequence.
ACCESSION CC803004
VERSION CC803004.1 GI:32627260
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 689)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
  FORWARD: T7
  BACKWARD: M13r
Plate: 0117 row: C column: 08
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
  1. .689
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /cultivar="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBBb0117C08"
  /lab_host="DH10B"
  /clone_lib="ZMMBBb"
  /note="Vector: pBeloBAC1.1; Site_1: HindIII; Site_2: HindIII; Zea mays L. ssp. mays"

ORIGIN
  Query Match      80.9%; Score 17.8; DB 9; Length 689;
  Best Local Similarity 90.5%; Pred. NO. 8.3e+02;
  Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 584 ACCAAGATGGTGCATCTGAT 564

RESULT 30
CC803004 690 bp DNA linear GSS 14-JUL-2003
LOCUS ih10b03.g1 WGS-ZmaysU (DH10b or JM107MA2) Zea mays genomic clone
DEFINITION ih10b03, genomic survey sequence.
ACCESSION CC803004
VERSION CC803004.1 GI:32627260
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 689)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
  FORWARD: T7
  BACKWARD: M13r
Plate: 0117 row: C column: 08
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
  1. .689
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /cultivar="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBBb0117C08"
  /lab_host="DH10B"
  /clone_lib="ZMMBBb"
  /note="Vector: pBeloBAC1.1; Site_1: HindIII; Site_2: HindIII; Zea mays L. ssp. mays"

ORIGIN
  Query Match      80.9%; Score 17.8; DB 9; Length 689;
  Best Local Similarity 90.5%; Pred. NO. 8.3e+02;
  Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 584 ACCAAGATGGTGCATCTGAT 564

RESULT 30
CC803004 690 bp DNA linear GSS 14-JUL-2003
LOCUS ih10b03.g1 WGS-ZmaysU (DH10b or JM107MA2) Zea mays genomic clone
DEFINITION ih10b03, genomic survey sequence.
ACCESSION CC803004
VERSION CC803004.1 GI:32627260
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 689)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
  FORWARD: T7
  BACKWARD: M13r
Plate: 0117 row: C column: 08
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
  1. .689
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /cultivar="B73"
  /db_xref="taxon:4577"
  /clone="ZMMBBb0117C08"
  /lab_host="DH10B"
  /clone_lib="ZMMBBb"
  /note="Vector: pBeloBAC1.1; Site_1: HindIII; Site_2: HindIII; Zea mays L. ssp. mays"

ORIGIN
  Query Match      80.9%; Score 17.8; DB 9; Length 689;
  Best Local Similarity 90.5%; Pred. NO. 8.3e+02;
  Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 584 ACCAAGATGGTGCATCTGAT 564

RESULT 30
CC803004 690 bp DNA linear GSS 14-JUL-2003
LOCUS ih10b03.g1 WGS-ZmaysU (DH10b or JM107MA2) Zea mays genomic clone
DEFINITION ih10b03, genomic survey sequence.
ACCESSION CC803004
VERSION CC803004.1 GI:32627260
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

1 (bases 1 to 690)
Rabinowicz, P.D., O'Shaughnessy, A.L., Palmer, L., Balija, V., Dedhia, N., Katzenburger, F., King, L., Miller, B., Muller, S., Nascento, L., Zutavern, T., McCombie, W.R. and Martienssen, R.A.
Genomic shotgun sequences from Zea mays (unfiltered)
Unpublished (2003)

TITLE

JOURNAL

COMMENT

Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: ih10 row: b column: 03

Seq primer: -21M13UnivRev

Class: shotgun

High quality sequence stop: 690.

FEATURES

source

1..690
Location/Qualifiers
/organism="Zea mays"
/mol_type="Genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ih10b03"
/lab_host="DH10b or JM107MA2"
/clone_lib="WGS-ZmaysU (DH10b or JM107MA2)"
/note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The ligation was transformed into DH10b or JM107MA2."

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 690;
Best Local Similarity 90.5%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21

||||| ||||| ||||| |||||

Db 185 ACCCAAGATGGTGTCATCTGAT 205

RESULT 31

BZ998359/c

LOCUS

DEFINITION
BZ998359 PUDPH91TD ZM 0.6.1.0 KB Zea mays genomic clone ZMMBTal74014,
genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

Clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 691)

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennettzen, J.

Maize Genomics Consortium

Unpublished (2003)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

Location/Qualifiers

1..691

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBTal74014"

/clone_lib="ZM 0.6.1.0 KB"

/note="Vector: PCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high

Cot selected genomic DNA library"

ORIGIN

Query Match 80.9%; Score 17.8; DB 8; Length 691;
Best Local Similarity 90.5%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21

||||| ||||| ||||| |||||

Db 181 ACCCAAGATGGTGTCATCTGAT 161

RESULT 32

CC629672/c

LOCUS

DEFINITION

CC629672 ZM 0.7.1.5 KB Zea mays genomic clone ZMMBma0460120,
genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Ze mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

Clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 693)

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSS: OGUKF58TV

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..693

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMMBma0460120"

/clone_lib="ZM 0.7.1.5 KB"

/note="Vector: pBCSK-; Site_1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"

ORIGIN

Query Match 80.9%; Score 17.8; DB 9; Length 693;
Best Local Similarity 90.5%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21

||||| ||||| ||||| |||||

Db 192 ACCCAAGATGGTGTCATCTGAT 172

RESULT 33

```

CC464996
LOCUS          CC464996          697 bp      DNA      linear      GSS 04-JUN-2003
DEFINITION    ZMMBBc0385B04r Zea mays genomic clone ZMMBBc0385B04 3',
genomic survey sequence.
ACCESSION     CC464996
VERSION       CC464996.1 GI:31382585
KEYWORDS      GSS.
SOURCE        Zea mays
ORGANISM      Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 697)
AUTHORS      Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
               Rourard,K., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE        Sequencing of the maize genome at PGIR (2003b)
JOURNAL       Unpublished (2003)
COMMENT      Contact: Bharti,A.K.
               Dr. Joachim Messing's lab
               The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
               University
               190 Frelinghuysen Road, Piscataway, NJ 08854, USA
               Tel: 732 445 3801
               Fax: 732 445 5735
               Email: bharti@waksman.rutgers.edu
               Seq primer: SP6
               Class: BAC ends
               High quality sequence start: 65.
               Location/Qualifiers
                 1..697
                   /organism="Zea mays"
                   /mol_type="genomic DNA"
                   /cultivar="B73"
                   /db_xref="taxon:4577"
                   /clone="ZMMBBc0385B04"
                   /lab_host="E. coli DH10B"
                   /clone_lib="ZMMBBc"
                   /notes="Vector: pTARBAC1.3; Site_1: BamHI; Site_2: BamHI"

FEATURES             source
  ORIGIN
    Query Match      80.9%; Score 17.8; DB 8; Length 697;
    Best Local Similarity 90.5%; Pred. No. 8.3e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 227 ACCCAAGATGGTGCATCTGAT 247

  RESULT 34
  CG913717/c
  LOCUS          CG913717        697 bp      DNA      linear      GSS 09-DEC-2003
  DEFINITION    ZMMBBb0373P04.f ZMMBBb Zea mays genomic clone ZMMBBb0373P04 5',
  genomic survey sequence.
  ACCESSION     CG913717
  VERSION       CG913717.1 GI:39613003
  KEYWORDS      GSS.
  SOURCE        Zea mays
  ORGANISM      Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
  REFERENCE     1 (bases 1 to 697)
  AUTHORS      Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
               and Wing,R.
  TITLE        Sequencing of the maize genome
  JOURNAL       Unpublished (2003)
  COMMENT      Contact: Rod Wing
               Arizona Genomics Institute
               University of Arizona
               Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
               85721-0088, USA
               Tel: 520 626 3967

```

```

Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
  FORWARD: T7
  BACKWARD: M13r
Plate: 0373 row: P column: 04
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
  1..697
    /organism="Zea mays"
    /mol_type="genomic DNA"
    /cultivar="B73"
    /db_xref="taxon:4577"
    /clone="ZMMBBb0373P04"
    /lab_host="DH10B"
    /clone_lib="ZMMBBb"
    /note="Vector: pBelOBAC11; Site 1: HindIII; Site 2:
    HindIII; Zea mays L. sep. mays"

FEATURES             source
  ORIGIN
    Query Match      80.9%; Score 17.8; DB 9; Length 697;
    Best Local Similarity 90.5%; Pred. No. 8.3e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

  QY 1 ACCCATGATGGCGCATCTGAT 21
      ||||| ||||| ||||| |||||
  Db 663 ACCCAAGATGGTGCATCTGAT 643

  RESULT 35
  CC159056/c
  LOCUS          CC159056        698 bp      DNA      linear      GSS 29-APR-2003
  DEFINITION    if98b03.b1 WGS-ZmayrF (DH5a methyl filtered) Zea mays genomic clone
  if98b03, genomic survey sequence.
  ACCESSION     CC159056
  VERSION       CC159056.1 GI:30183834
  KEYWORDS      GSS.
  SOURCE        Zea mays
  ORGANISM      Zea mays
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
               clade; Panicoideae; Andropogoneae; Zea.
  REFERENCE     1 (bases 1 to 698)
  AUTHORS      Rabinowicz,P.D., O'Shaughnessy,A.L., Balijs,V., Dedhia,N.,
               Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
               Zutavern,T., McCombie,W.R. and Martienssen,R.A.
               Genomic shotgun sequences from Zea mays (methyl-filtered)
               Unpublished (2002)
               Contact: W. Richard McCombie
               Lita Annenberg Hazen Genome Sequencing Center
               Cold Spring Harbor Laboratory
               PO Box 100, Cold Spring Harbor, NY 11724, USA
               Tel: 516 367 8884
               Fax: 516 367 8874
               Email: mcombie@cshl.org
               Plate: if98 row: b column: 03
               Seq primer: -21M13UnivFwd
               Class: shotgun
               High quality sequence stop: 698.
               Location/Qualifiers
                 1..698
                   /organism="Zea mays"
                   /mol_type="genomic DNA"
                   /cultivar="B73"
                   /db_xref="taxon:4577"
                   /clone="if98b03"
                   /lab_host="DH5a"
                   /clone_lib="WGS-ZmayrF (DH5a methyl filtered)"
                   /note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
                   The vector was digested with XbaI and one nucleotide was
                   added by fill in in the recessive 3' end. The genomic DNA
                   was nebulized, end repaired, adaptor ligated and size

```

fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed into DH5a."

```

ORIGIN
Query Match      80.9%; Score 17.8; DB 8; Length 698;
Best Local Similarity 90.5%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 211 ACCCAAGATGGTGCACTCGAT 191

RESULT 36
CC744522/c
LOCUS CC744522/c
DEFINITION ZMMBBb0120A13.f ZMMBBb Zea mays genomic clone ZMMBBb0120A13 5',
genomic survey sequence.
ACCESSION CC744522
VERSION CC744522
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 703)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0120 row: A column: 13
Seq primer: T7
Class: BAC ends.
FEATURES
source
1..703
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBb0120A13"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. sep. mays"

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 703;
Best Local Similarity 90.5%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 483 ACCCAAGATGGTGCACTCGAT 463

RESULT 37
CC788457/c
LOCUS CC788457/c
DEFINITION ZMMBBb0156P21.f ZMMBBb Zea mays genomic clone ZMMBBb0156P21 5',
genomic survey sequence.
ACCESSION CG352166
VERSION CG352166
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 708)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGVGB28TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

```

```

ACCESSION CC788457
VERSION CC788457.1 GI:32383680
KEYWORDS GSS
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 704)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J.
and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: T7
BACKWARD: M13r
Plate: 0156 row: P column: 21
Seq primer: T7
Class: BAC ends.
FEATURES
source
1..704
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBBb0156P21"
/lab_host="DH10B"
/clone_lib="ZMMBBb"
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2:
HindIII; Zea mays L. sep. mays"

ORIGIN
Query Match      80.9%; Score 17.8; DB 9; Length 704;
Best Local Similarity 90.5%; Pred. No. 8.3e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
||||| ||||| ||||| ||||| |||||
Db 677 ACCCAAGATGGTGCACTCGAT 657

RESULT 38
CG352166/c
LOCUS CG352166/c
DEFINITION OGVGB28TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0488E08,
genomic survey sequence.
ACCESSION CG352166
VERSION CG352166.1 GI:34269432
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 708)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: OGVGB28TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA

```



```
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
    source
        1..708
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMWBa048808"
            /clone_lib="ZM 0.7-1.5_KB"
            /notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"

ORIGIN
    Query Match      80.9%; Score 17.8; DB 9; Length 708;
    Best Local Similarity 90.5%; Pred. No. 8.3e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 605 ACCCAAGATGGTGCATCTGAT 585

RESULT 39
CC876585/c
LOCUS
DEFINITION
    ZMWBb0195G21.f ZMWBb Zea mays genomic clone ZMWBb0195G21 5'.
    genomic survey sequence.
ACCESSION
    CC876585
VERSION
    CC876585.1 GI:33306698
KEYWORDS
    GSS.
SOURCE
    Zea mays
    ORGANISM
        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
        Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
        clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
    1 (bases 1 to 711)
    Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.
    and Wing, R.
    Sequencing of the maize genome
    Unpublished (2003)
    Contact: Rod Wing
    Arizona Genomics Institute
    University of Arizona
    Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
    85721-0088, USA
    Tel: 520 626 3967
    Fax: 520 621 9288
    Email: http://genome.arizona.edu
    PCR Primers
    FORWARD: T7
    BACKWARD: M13r
    Plate: 0195 row: G column: 21
    Seq primer: T7
    Class: BAC ends.
FEATURES
    source
        1..711
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /cultivar="B73"
            /db_xref="taxon:4577"
            /clone="ZMWBb0195G21"
            /lab_host="DH10B"
            /clone_lib="ZMWBb"
            /notes="Vector: pBelOBAC11; Site 1: HindIII; Site 2:
            HindIII; Zea mays L. esp. mays"

ORIGIN
    Query Match      80.9%; Score 17.8; DB 9; Length 711;
    Best Local Similarity 90.5%; Pred. No. 8.3e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 595 ACCCAAGATGGTGCATCTGAT 575

Search completed: June 4, 2005, 11:46:43
Job time : 1346.17 secs
```

```
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
    source
        1..708
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMWBa048808"
            /clone_lib="ZM 0.7-1.5_KB"
            /notes="Vector: pBCK-; Site 1: HincII; 0.7-1.5 kb
            methylation filtered genomic DNA library"

ORIGIN
    Query Match      80.9%; Score 17.8; DB 9; Length 708;
    Best Local Similarity 90.5%; Pred. No. 8.3e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 605 ACCCAAGATGGTGCATCTGAT 585

RESULT 39
CC876585/c
LOCUS
DEFINITION
    ZMWBb0195G21.f ZMWBb Zea mays genomic clone ZMWBb0195G21 5'.
    genomic survey sequence.
ACCESSION
    CC876585
VERSION
    CC876585.1 GI:33306698
KEYWORDS
    GSS.
SOURCE
    Zea mays
    ORGANISM
        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
        Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
        clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
    1 (bases 1 to 711)
    Yu, Y., Kim, H.R., Hatfield, J., Soderlund, C., Bharti, A.K., Messing, J.
    and Wing, R.
    Sequencing of the maize genome
    Unpublished (2003)
    Contact: Rod Wing
    Arizona Genomics Institute
    University of Arizona
    Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
    85721-0088, USA
    Tel: 520 626 3967
    Fax: 520 621 9288
    Email: http://genome.arizona.edu
    PCR Primers
    FORWARD: T7
    BACKWARD: M13r
    Plate: 0195 row: G column: 21
    Seq primer: T7
    Class: BAC ends.
FEATURES
    source
        1..711
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /cultivar="B73"
            /db_xref="taxon:4577"
            /clone="ZMWBb0195G21"
            /lab_host="DH10B"
            /clone_lib="ZMWBb"
            /notes="Vector: pBelOBAC11; Site 1: HindIII; Site 2:
            HindIII; Zea mays L. esp. mays"

ORIGIN
    Query Match      80.9%; Score 17.8; DB 9; Length 711;
    Best Local Similarity 90.5%; Pred. No. 8.3e+02;
    Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCCATGATGGCGCATCTGAT 21
    ||||| ||||| ||||| |||||
Db 595 ACCCAAGATGGTGCATCTGAT 575

Search completed: June 4, 2005, 11:46:43
Job time : 1346.17 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 04:22:34 ; Search time 234.976 Seconds
(without alignments)
780.981 Million cell updates/sec

Title: US-09-674-277-25
Perfect score: 31
Sequence: 1 acgttctgtcttaacggtgatgtaggtttt 31

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	31	Aaz36125	Aaz36125 Primer de
2	31	100.0	1181	Aaz36102	Aaz36102 Nucleic a
3	21	67.7	411	Aah94232	Aah94232 Human foe
4	21	67.7	982	Aah94566	Aah94566 Human foe
5	20.4	65.8	1023	Adf02519	Adf02519 Bacterial
6	20	64.5	1098	AcF74971	AcF74971 Staphyloc
7	19.8	63.9	1425	Ab103085	Ab103085 Drosophil
8	19.8	63.9	9466	Ab103084	Ab103084 Drosophil
9	19.8	63.9	10026	Ab133876	Ab133876 Human imm
10	19.8	63.9	107543	Abd33524	Abd33524 Human can
11	19.6	63.2	1578	Abn70175	Abn70175 Streptoco
12	19.6	63.2	1581	ACA50695	ACA50695 Prokaryot
13	19.4	62.6	576	ABV01466	ABV01466 Human pro
14	19.4	62.6	6129	AB134582	AB134582 Human met
15	19.4	62.6	6129	Ades99843	Ades99843 Bisulphit
16	19.2	61.9	162	ACA30142	ACA30142 Prokaryot
17	19.2	61.3	1535	ABL29753	ABL29753 Drosophil
18	19	61.3	2853	ABd13837	ABd13837 Pseudomon
19	19	61.3	3864	ABD14431	ABD14431 Pseudomon
20	19	61.3	4629	ACC50111	ACC50111 Breast ca

	21	19	61.3	4629	9	ADB70381	ADB70381 Procollag
	22	19	61.3	4629	10	ADD15214	ADD15214 Human NRG
	23	19	61.3	4629	11	ADP65599	ADP65599 Human mRN
	24	19	61.3	4629	12	ADJ37160	ADJ37160 Human mal
	25	19	61.3	4629	13	ADR99115	ADR99115 Collagen,
C	26	19	61.3	6134	4	ABL29752	ABL29752 Drosophil
	27	19	61.3	6217	11	ADP65055	ADP65055 Human col
	28	19	61.3	6217	11	ADP65717	ADP65717 Human col
C	29	19	61.3	6792	4	ABL29978	ABL29978 Drosophil
	30	18.8	60.6	211	5	ABV31805	ABV31805 Human pro
	31	18.8	60.6	407	5	ABV12481	ABV12481 Human pro
	32	18.8	60.6	444	5	ABV03312	ABV03312 Human pro
	33	18.8	60.6	446	5	ABV42537	ABV42537 Human pro
	34	18.8	60.6	446	5	ABV33617	ABV33617 Human pro
	35	18.8	60.6	446	5	ABV42534	ABV42534 Human pro
	36	18.8	60.6	446	5	ABV40772	ABV40772 Human pro
	37	18.8	60.6	462	8	ACA52379	ACA52379 Prokaryot
C	38	18.8	60.6	489	6	ABN22910	ABN22910 Human ORF
	39	18.8	60.6	496	5	ABV10635	ABV10635 Human pro
	40	18.8	60.6	751	6	ABQ38090	ABQ38090 Oligonucle
C	41	18.8	60.6	751	6	ABQ38091	ABQ38091 Oligonucle
	42	18.8	60.6	897	5	ABV22160	ABV22160 Human pro
	43	18.8	60.6	897	5	ABV23213	ABV23213 Human pro
	44	18.8	60.6	897	5	ABV29053	ABV29053 Human pro
	45	18.8	60.6	897	5	ABV26782	ABV26782 Human pro

ALIGNMENTS

RESULT 1

Aaz36125

ID Aaz36125 standard; DNA; 31 BP.

AC Aaz36125;

DT 11-FEB-2000 (first entry)

DE Primer derived from a nucleic acid sequence specific to EHEC.

KW Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virk gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

XX WO9955908-A2.

PD 04-NOV-1999.

PF 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PI Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic Escherichia coli, particularly serotype O157:H7, used for detecting these bacteria in food.

XX Claim 5; Page 27; 48pp; French.

XX AAZ36103-27 represent fragments derived from nucleic acid sequences specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are derived from two sequences. The first (AAZ36101) is 99.9% homologous to the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102). The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
 CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
 CC for virulence proteins of Shigella flexneri. Both sequences are of
 CC plasmid origin. The fragments are used as PCR primers and probes for the
 CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
 CC in human or animal samples, foods or the environment. The fragments are
 CC also useful for epidemiological studies

XX
 SQ Sequence 31 BP; 4 A; 4 C; 10 G; 13 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGTTCGTCTTACCGGGTGATGAGTTT 31
 |||||
 Db 1 ACGTTCGTCTTACCGGGTGATGAGTTT 31

RESULT 2

AAZ36102
 ID AAZ36102 standard; DNA; 1181 BP.

XX
 AC AAZ36102;

XX
 DT 11-FEB-2000 (first entry)

XX Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

DE Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

XX enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.

XX Escherichia coli.

XX WO955908-A2.

XX 04-NOV-1999.

XX 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
 PT Escherichia coli, particularly serotype O157:H7, used for detecting these
 PT bacteria in food.

XX Claim 1; Fig 2; 48pp; French.

XX The present sequence is specific to enterohemorrhagic Escherichia coli
 CC (EHEC). The sequence associated with the presence of virulence factors
 CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
 CC 68% homology with the virK gene which codes for virulence proteins of
 CC Shigella flexneri. The present sequence is of plasmid origin. Fragments
 CC of the present sequence are used, as probes and primers, for detection of
 CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
 CC animal samples, foods or the environment. The fragments are also useful
 CC for epidemiological studies

SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 1181;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACGTTCGTCTTACCGGGTGATGAGTTT 31
 |||||
 Db 928 ACGTTCGTCTTACCGGGTGATGAGTTT 958

RESULT 3

AAH94232

ID AAH94232 standard; cDNA; 411 BP.

XX
 AC AAH94232;

XX 05-OCT-2001 (first entry)

XX Human foetal cDNA, SEQ ID NO: 761.

XX Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
 KW nontropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
 KW gene therapy; antisense therapy; cancer; immune disorder;
 KW growth disorder; osteoporosis; thrombolytic disorder;
 KW nervous system disorder; inflammation; expressed sequence tag; EST; ss.

XX Homo sapiens.

XX WO200155339-A2.

XX 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US002723.

XX 25-JAN-2000; 2000US-00491404.

XX 15-SEP-2000; 2000US-00663870.

XX 06-NOV-2000; 2000US-00707351.

XX (HYSE-) HYSEQ INC.

XX Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;

XX Liu C, Asundi V, Zhou P, Werhman T;

XX WPI; 2001-465571/50.

XX P-PSDB; AAM06557.

XX Novel fetal proteins useful for the treatment and diagnosis of diseases

XX associated with dysfunction of the protein e.g. cancers, immune

XX disorders, growth disorders, thrombolytic disorders, nervous system

XX disorders and inflammation.

XX Claim 1; Page 466; 715pp; English.

XX The invention relates to novel foetal polypeptides encoded by
 CC polynucleotides comprising one of 477 sequences fully defined in the
 CC specification. The foetal polynucleotides and polypeptides are useful in
 CC the treatment and diagnosis of diseases such as cancers, immune

CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,

CC nervous system disorders and inflammation. The present sequence was

CC assembled using an expressed sequence tag (EST) found to be expressed in

CC human foetal tissue cDNA libraries as the seed

SQ Sequence 411 BP; 103 A; 65 C; 80 G; 161 T; 0 U; 2 Other;

Query Match 67.7%; Score 21; DB 5; Length 411;
 Best Local Similarity 82.8%; Pred. No. 39;
 Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 GTTCTGGTCTTACCGGGTGATGAGTTT 31
 |||||
 Db 14 GTTCTGGTCTTACCGGGTGATGAGTTT 42

RESULT 4

AAH94566

ID AAH94566 standard; cDNA; 982 BP.

XX
 AC AAH94566;

XX 05-OCT-2001 (first entry)

XX


```
RESULT 9
ABL33876
ID ABL33876 standard; DNA; 10026 BP.
XX
AC ABL33876;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1849.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antineumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPITG-) EPITENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
WPI; 2002-130909/17.
XX
KW Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 1849; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
XX
SQ Sequence 10026 BP; 2081 A; 337 C; 2786 G; 4822 T; 0 U; 0 Other;
XX
Query Match 63.9%; Score 19.8; DB 6; Length 10026;
Best Local Similarity 77.4%; Pred. No. 2e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
QY 1 ACGTTCTGCTTACGGGTGATGAGTTT 31
| | | | | | | | | | | | | | | | | | | | | |
DB 3252 AAGTTTGTGCTGATGGGTGAGTAGTTT 3282
| | | | | | | | | | | | | | | | | | | | | |
XX
RESULT 10
ABD33524
ID ABD33524 standard; DNA; 107543 BP.
XX
AC ABD33524;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human cancer-associated (CA) gene HD07-103.
XX
KW Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
ds; cancer; cytostatic.
KW Homo sapiens.
WO2004058146-A2.
15-JUL-2004.
15-DEC-2003; 2003WO-US040081.
17-DEC-2002; 2002US-00322281.
(SAGR-) SAGRES DISCOVERY INC.
Morris DW, Melandro MS;
WPI; 2004-499109/47.
Novel human cancer associated protein encoded within open reading frame
of cancer associated gene, useful as targets for diagnosing cancer.
Claim 16; SEQ ID NO 706; 182pp; English.
The invention relates to cancer-associated proteins (CAP) and the cancer-
associated (CA) nucleic acids encoding them. The invention also relates
to a method for treating cancers involving administering to a patient an
inhibitor of CAP, and a method of screening for anticancer activity in a
potential drug involving providing a cell that expresses a CA gene,
contacting a tissue sample derived from a cancer cell with an anticancer
drug candidate and monitoring the effect of the anticancer drug candidate
on expression of the CA gene. The CAP proteins are useful for detecting
cancer associated with expression of a CAP protein in a test cell sample
and for screening for a bioactive agent capable of modulating the
activity of a CAP protein. The CA nucleic acids are useful for diagnosing
cancer, involving determining the expression of a CA nucleic acid in a
tissue. This sequence represents a human CA gene of the invention. Note:
The sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 107543 BP; 30009 A; 20779 C; 22714 G; 34041 T; 0 U; 0 Other;
XX
Query Match 63.9%; Score 19.8; DB 13; Length 107543;
Best Local Similarity 77.4%; Pred. No. 2.7e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
QY 1 ACGTTCTGCTTACGGGTGATGAGTTT 31
| | | | | | | | | | | | | | | | | | | | | |
DB 65009 ATGGTTTGTGTTACGGGTGATGTTGTTCT 65039
| | | | | | | | | | | | | | | | | | | | | |
XX
RESULT 11
ABN70175
ID ABN70175 standard; DNA; 1578 BP.
XX
AC ABN70175;
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 8263.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB004789.
```

```
PR 27-OCT-2000; 2000GB-00026333.
PR 24-NOV-2000; 2000GB-00028727.
PR 07-MAR-2001; 2001GB-00005640.
XX
XX (CHIR-) CHIRON SPA.
XX (GENO-) INST GENOMIC RES.
XX
PI Telford J, Maignani V, Margarit Y Rosl, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI: 2002-352536/38.
DR P-PSDB; ABP29544.
XX
XX New Streptococcus protein for the treatment or prevention of infection or
PT disease caused by Streptococcus bacteria, such as meningitis, and for
PT detecting a compound that binds to the protein.
XX
XX Claim 7; Page 3944; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins
XX
XX Sequence 1578 BP; 536 A; 312 C; 272 G; 458 T; 0 U; 0 Other;
SQ
Query Match 63.2%; Score 19.6; DB 6; Length 1578;
Best Local Similarity 84.6%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 TTCTGGTCTTACGGGTGATGTAGGTT 29
Db 243 TTCTGTTGTTACTGCTGATTAGGTT 268
RESULT 12
ACA50695
ID ACA50695 standard; DNA; 1581 BP.
AC
AC ACA50695;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #32352.
XX
KW Antisense; ds: prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Streptococcus pyogenes.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR
PR 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI: 2003-029926/02.
DR P-PSDB; ABU46825.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 38565; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway; (8)
CC required for proliferation, or that inhibits cellular proliferation; (9)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1581 BP; 538 A; 312 C; 272 G; 459 T; 0 U; 0 Other;
SQ
Query Match 63.2%; Score 19.6; DB 8; Length 1581;
Best Local Similarity 84.6%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 TTCTGGTCTTACGGGTGATGTAGGTT 29
Db 243 TTCTGTTGTTACTGCTGATTAGGTT 268
RESULT 13
ABV01466
ID ABV01466 standard; cDNA; 576 BP.
XX
XX AC ABV01466;
XX
XX 13-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 1457.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
```


XX WO200160860-A2.
PN
XX
XX PD
XX
XX PF
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 324; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker .
XX
XX Sequence 576 BP; 138 A; 121 C; 113 G; 153 T; 0 U; 51 Other;
DE
XX Query Match 62.6%; Score 19.4; DB 5; Length 576;
Best Local Similarity 76.7%; Pred. No. 2e+02; Mismatches 0; Gaps 0;
Matches 23; Conservative 0; Indels 7; Indels 0; Gaps 0;
QY 1 ACGTTCGCTCTTACGGGTGATGTAGGTTT 30
Db 20 ACGTTCGGAATTCGGGAGATCGGNTTT 49
RESULT 14
ABL34582
ID ABL34582 standard; DNA; 6129 BP.
XX
XX ABL34582;
AC
XX 26-MAR-2002 (first entry)
DT
XX
XX Human metastasis associated gene SEQ ID NO: 135.
DE
XX Metastasis associated gene; cytostatic; gene therapy; cancer;
KW Cytosine methylation; gene; ds.
XX
XX Homo sapiens.
OS
XX WO200177376-A2.
PN
XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-EP003970.
PF
XX
XX 06-APR-2000; 2000DE-01019058.
PR
XX 07-APR-2000; 2000DE-01019173.
PR

PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-010922/01.
DR
XX New nucleic acid derived from chemically treated metastasis genes, useful
PT for diagnosis of cancers by analysis of cytosine methylation, also for
PT treatment.
XX
XX Claim 1; SEQ ID NO 135; 23pp + Sequence Listing; English.
PS
XX The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention
XX
XX Sequence 6129 BP; 1620 A; 79 C; 1538 G; 2892 T; 0 U; 0 Other;
SQ
XX Query Match 62.6%; Score 19.4; DB 6; Length 6129;
Best Local Similarity 79.3%; Pred. No. 2.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 GTTCTGGTCTTACGGGTGATGTAGGTTT 31
Db 1181 GTTGTGTGTTTCGGGGAGGTAGGTTT 1209
RESULT 15
ADS99843
ID ADS99843 standard; DNA; 6129 BP.
XX
XX ADS99843;
AC
XX 02-DEC-2004 (first entry)
DT
XX
XX Bisulphite treated human gene associated with metastasis #68.
DE
XX Human; ds; Gene; Bisulphite; metastasis; cancer; cytostatic;
KW DNA methylation; matrix-assisted laser desorption/ionisation; MALDI;
KW electrospray; mass spectrometry; CpG dinucleotide; solid tumour.
XX
XX Homo sapiens.
OS
XX US2003148327-A1.
PN
XX
XX 07-AUG-2003.
PD
XX
XX 21-JAN-2003; 2003US-00240485.
PF
XX
XX 06-APR-2000; 2000DE-01019058.
PR
XX 07-APR-2000; 2000DE-01019173.
PR
XX 30-JUN-2000; 2000DE-01032529.
PR
XX 01-SEP-2000; 2000DE-01043826.
PR
XX 06-APR-2001; 2001WO-EP003970.
PR
XX (OLEK/) OLEK A.
PA (PIEP/) PIEPENBROCK C.
PA (BERL/) BERLIN K.
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-010922/01.
DR
XX New nucleic acid derived from chemically treated metastasis genes, useful
PT for diagnosis of cancers by analysis of cytosine methylation, also for
PT treatment.
XX
XX Claim 1; SEQ ID NO 135; 9pp; English.
PS
XX

PS Disclosure; SEQ ID NO 13035; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

CC polynucleotides encoding them. The sequences are useful in diagnosis and

CC therapy of pathological conditions, as molecular targets for diagnostics,

CC prophylaxis and treatment of pathological conditions resulting from a

CC bacterial infection, for evaluating a compound, such as a polypeptide, of

CC for the ability to bind a P. aeruginosa nucleic acid, as components of

CC effective antibacterial targets, as targets for antibacterial drugs,

CC including anti-P. aeruginosa drugs, as templates for recombinant

CC production of P. aeruginosa-derived peptides or polypeptides, as target

CC components for diagnosis and/or treatment of P. aeruginosa-caused

CC infection, and in detection of P. aeruginosa sequences or other sequences

CC of Pseudomonas species using biochip technology. Sequences ABD01397-

CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:

CC The sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format from USPTO at

CC seqdata.uspto.gov/sequence.html

XX

XX Sequence 3864 BP; 773 A; 1130 C; 1268 G; 693 T; 0 U; 0 Other;

XX

Query Match 61.3%; Score 19; DB 11; Length 3864;

Best Local Similarity 81.5%; Pred. No. 3.8e+02;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CGTTCGTCTTACCGGTGATGTAGGT 28

DB 2358 CGCGCGGTCTTACCGGTGATGTGGT 2332

RESULT 20

ACC50111

ID ACC50111 standard; cDNA; 4629 BP.

XX

AC ACC50111;

XX

DT 12-JUN-2003 (first entry)

XX

DE Breast cancer associated cDNA sequence SEQ ID NO:69.

XX

Human; breast cancer; cytostatic; gene therapy; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003004989-A2.

XX

PD 16-JAN-2003.

XX

PF 21-JUN-2002; 2002WO-US019669.

XX

PR 21-JUN-2001; 2001US-0299887P.

PR 27-JUN-2001; 2001US-0301572P.

PR 18-JUL-2001; 2001US-0306501P.

PR 25-SEP-2001; 2001US-0325002P.

PR 05-MAR-2002; 2002US-0362585P.

PR 14-MAY-2002; 2002US-0380391P.

XX

(MILL-) MILLENIUM PHARM INC.

XX

Lillie J, Gannavarapu M, Glatt K, Hoersht S, Kanatkar S;

PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;

PI East RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX

WPI; 2003-210381/20.

DR P-PSDB; ABR47419.

XX

Breast cancer diagnosis or treatment by comparing the level of expression

PT of a marker in a patient sample with that in the control non-breast

PT cancer sample.

XX

Claim 1; SEQ ID NO 69; 128pp; English.

PS

XX The present invention describes a method for assessing whether a patient

CC

CC is afflicted with breast cancer. The method comprises comparing the level

CC of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and

CC ABR47386 to ABR47632) in a patient sample and the normal level of

CC expression of the marker in a control non-breast cancer sample, where a

CC significant increase in the level of expression of the marker in the

CC patient sample and the normal level is an indication that the patient is

CC afflicted with breast cancer. The breast cancer associated sequences from

CC the present invention have cytostatic activities and can be used in gene

CC therapy. The method is useful for diagnosing and treating breast cancer.

CC N.B. The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences

XX

XX Sequence 4629 BP; 1130 A; 1127 C; 1454 G; 918 T; 0 U; 0 Other;

XX

Query Match 61.3%; Score 19; DB 8; Length 4629;

Best Local Similarity 81.5%; Pred. No. 3.9e+02;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACCGGTGATGTAGGTTT 30

DB 1527 TTGGGCTTCCGGTGATCCAGGTTT 1553

RESULT 21

ADB70381

ID ADB70381 standard; cDNA; 4629 BP.

XX

AC ADB70381;

XX

DT 04-DEC-2003 (first entry)

XX

DE Procollagen alpha 2(V) cDNA SEQ ID NO:73.

XX

KW cancer; malignant pleural mesothelioma; MPM; lung adenocarcinoma;

KW squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;

KW diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;

KW human; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003021229-A2.

XX

PD 13-MAR-2003.

XX

PF 05-SEP-2002; 2002WO-US028203.

XX

PR 05-SEP-2001; 2001US-0317389P.

PR 30-AUG-2002; 2002US-00236031.

XX

(BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX

Gordon GJ, Jensen RV, Gullans SR, Bueno R;

PI WPI; 2003-290233/28.

DR P-PSDB; ADB70382.

XX

Diagnosing cancer cells in tissue sample, or determining prognosis or

PT outcome of cancer patient, by calculating ratio of expression levels of

PT genes that are differentially expressed in cancer and non cancer tissues.

XX

Claim 67; Page 382-384; 396pp; English.

XX

The present invention describes a method (M1) for diagnosing the presence

CC of cancer cells or non-cancer cells in a tissue sample, or determining

CC the prognosis or outcome of a cancer patient. M1 involves providing a set

CC of genes that are differentially expressed in cancerous or non-cancerous

CC conditions, determining the expression levels of the set of genes and

CC calculating a ratio of the expression levels of the differentially

CC expressed genes. M1 is useful for diagnosing the presence of cancer cells

CC or non-cancer cells in a tissue sample, where the cancer is malignant

CC pleural mesothelioma (MPM), lung adenocarcinoma, squamous carcinoma,

CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell

CC methods and compositions of the present invention are useful for
 CC diagnosing and treating autoimmune disease or arthritides, such as
 CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
 CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
 CC immune disease caused by an infectious agent. This polynucleotide
 CC represents a DNA sequence relating to the genes used in the analysis and
 CC treatment of autoimmune diseases or arthritides. Note: This sequence is
 CC not shown in the specification. It has been supplied in an electronic
 CC format from WIPO.
 XX
 SQ Sequence 4629 BP; 1130 A; 1127 C; 1454 G; 918 T; 0 U; 0 Other;

Query Match 61.3%; Score 19; DB 11; Length 4629;
 Best Local Similarity 81.5%; Pred. No. 3.9e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGATGAGTTT 30
 |||||
 Db 1527 TTTCGGCCCTTCGGGTGATCCAGGTTT 1553

RESULT 24
 ADJ37160
 ID ADJ37160 standard; cDNA; 4629 BP.
 XX
 AC ADJ37160;

DT 22-APR-2004 (first entry)

DE Human malignant pleural mesothelioma (MPM) cDNA #33.

XX Human; malignant pleural mesothelioma; MPM; gene; ss; tumour;
 KW lung adenocarcinoma; squamous carcinoma; medulloblastoma;
 KW prostate cancer; breast cancer; diffuse large B-cell lymphoma;
 KW follicular lymphoma; ovarian cancer; cytostatic.

XX Homo sapiens.

XX US2003219760-A1.

XX 27-NOV-2003.

XX 05-SEP-2002; 2002US-00236031.

XX 05-SEP-2001; 2001US-0317389P.

PR 30-AUG-2002; 2002US-0407431P.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

XX Gordon GJ, Jensen RV, Gullans SR, Bueno R;

PI P-PSDB; ADJ37161.

DR WPI; 2004-141744/14.
 DR P-PSDB; ADJ37161.
 XX
 XX Diagnosing the presence of cancer or non-cancer cells in tissue sample,
 PT useful for diagnosing malignant pleural mesothelioma comprises
 PT determining ratio of expression level of a set of genes expressed in
 PT cancer tissues.

XX Claim 44; SEQ ID NO 73; 53pp; English.

XX The invention relates to a method of diagnosing the presence of cancer
 CC cells or non-cancer cells in a tissue sample, determining prognosis or
 CC outcome of a cancer patient, selecting a course of treatment for a
 CC subject having or suspected of having malignant pleural mesothelioma
 CC (MPM) and evaluating treatment of MPM comprising determining the ratio of
 CC the expression level of a set of genes differentially expressed in a
 CC cancer tissue. The cancer is chosen from MPM, lung adenocarcinoma,
 CC squamous carcinoma, medulloblastoma, prostate cancer, breast cancer,
 CC diffuse large B-cell lymphoma, follicular lymphoma and ovarian cancer.
 CC The method is useful for diagnosing MPM in a subject suspected of having
 CC MPM which involves obtaining a tissue sample suspected of being cancerous
 CC from a subject and determining the expression of nucleic acid markers or

CC its expression products in the tissue sample. This sequence represents
 CC human MPM cDNA of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification but was obtained in
 CC electronic format directly from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 4629 BP; 1130 A; 1127 C; 1454 G; 918 T; 0 U; 0 Other;

Query Match 61.3%; Score 19; DB 12; Length 4629;
 Best Local Similarity 81.5%; Pred. No. 3.9e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGATGAGTTT 30
 |||||
 Db 1527 TTTCGGCCCTTCGGGTGATCCAGGTTT 1553

RESULT 25
 ADR99115
 ID ADR99115 standard; DNA; 4629 BP.
 XX
 AC ADR99115;

DT 02-DEC-2004 (first entry)

DE Collagen, type V, alpha 2, COL5A2, coding sequence, SEQ ID 121.

XX Cytostatic; breast cancer; cancer; human; gene; ds;
 KW Collagen type V alpha 2; COL5A2.

XX Homo sapiens.

XX WO2004078035-A2.

XX 16-SEP-2004.

XX 27-FEB-2004; 2004WO-US0007268.

XX 28-FEB-2003; 2003US-0450655P.

XX (FARB) BAYER PHARM CORP.

XX Eveleigh D, Bigwood D;

XX WPI; 2004-653556/63.

DR P-PSDB; ADR99242.

DR REFSEQ; NM_000393.1.

XX Diagnosing breast cancer comprises comparing the level of expression of
 PT genes or gene products in a first biological sample taken from a patient
 PT with that in a normal patient sample.
 XX Claim 2; SEQ ID NO 121; 53pp; English.

XX The present invention relates to a method (M1) for diagnosing breast
 CC cancer in a patient. The method comprises comparing the level of
 CC expression of one or more genes or gene products in a biological sample
 CC from the patient with that in a normal patient sample, where a difference
 CC in the gene expression in the first sample compared to that in the second
 CC sample is a diagnostic of the disease. Also claimed are: method (M2) for
 CC distinguishing between normal and disease tissues; method (M3) for
 CC monitoring the response of a breast cancer patient to treatment with an
 CC anti-cancer agent; method (M4) for identifying a compound for treating
 CC breast cancer; and an array for distinguishing between normal and disease
 CC tissues comprising two or more probes corresponding to genes selected
 CC from ADR98995-ADR99121 or comprising two or more polypeptides selected
 CC from ADR99122-ADR99248. In M1 and M2 the genes are selected from ADR98995
 CC -ADR99121 and the gene products are polypeptides selected from ADR99122-
 CC ADR99248. M1 is useful for diagnosing breast cancer. M2 and the array are
 CC useful for distinguishing between normal and disease tissue. M3 is useful
 CC for monitoring the response of a breast cancer patient to treatment with
 CC an anti-cancer agent. M4 is useful for identifying a compound for
 CC treating breast cancer. Note: The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic

```
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4629 BP; 1130 A; 1127 C; 1454 G; 918 T; 0 U; 0 Other;

  Query Match      61.3%; Score 19; DB 13; Length 4629;
  Best Local Similarity 81.5%; Pred. No. 3.9e+02;
  Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGGTCTTACGGGTGATGATGTTT 30
Db 1527 TTCGGGCTTCCGGGTGATCAGGTTT 1553

RESULT 26
ABL29752/c
ID ABL29752 standard; DNA; 6134 BP.
XX
AC ABL29752;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 40729.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
KW Drosophila melanogaster.
XX
OS WO200171042-A2.
XX
PN 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US0009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX
XX Claim 1; SEQ ID NO 40729; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 6134 BP; 1897 A; 1246 C; 1196 G; 1795 T; 0 U; 0 Other;

  Query Match      61.3%; Score 19; DB 4; Length 6134;
  Best Local Similarity 81.5%; Pred. No. 4.1e+02;
  Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 TCTGGTCTTACGGGTGATGATGTTT 31
Db 1248 TCTGGTCTTCCGGGTGTTGTGTTGT 1222

RESULT 27
ADP65055
```

```
ID ADP65055 standard; DNA; 6217 BP.
XX
AC ADP65055;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human collagen, type V, alpha 2 (COL5A2) DNA sequence.
XX
KW autoimmune disease; arthritis; arthritide; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; antigout; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX
OS Homo sapiens.
XX
PN WO2003072827-A1.
XX
PD 04-SEP-2003.
XX
PF 31-OCT-2002; 2002WO-US035433.
XX
PR 31-OCT-2001; 2001US-0336220P.
XX
PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
PI Hirsch R, Thorton SL;
XX
XX WPI; 2003-712740/67.
XX
DR GENBANK; NM_000393.
XX
XX Diagnosing and analyzing autoimmune disease using gene expression
XX profiles and microarray technology, useful for diagnosing and treating
XX rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
XX gout.
XX
XX Disclosure; Page; 56pp; English.
XX
XX The invention relates to a novel method for diagnosing and analyzing
XX autoimmune disease or arthritides. The method comprises obtaining a
XX patient sample containing mRNA, analysing gene expression using the mRNA
XX that results in a gene expression signature of the mRNA, and using that
XX gene expression signature to diagnose or analyse the autoimmune disease
XX or arthritides in the patient, where gene expression of at least 60% of
XX the genes correlates with that of the gene signature. The invention
XX further comprises: a treatment of rheumatoid arthritis; identification of
XX genes for targeting in the treatment of rheumatoid arthritis in a mammal
XX other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
XX array or gene chip, specific for rheumatoid arthritis; diagnosis or
XX analyses of autoimmune disease or rheumatoid arthritis; screening the
XX efficacy of a candidate drug in vitro for the treatment of collagen-
XX induced arthritis; and reducing the symptoms associated with collagen-
XX induced arthritis. The compositions of the invention have the following
XX activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
XX antigout, antiinflammatory, dermatological, and immunomodulatory. The
XX methods and compositions of the present invention are useful for
XX diagnosing and treating autoimmune disease or arthritides, such as
XX rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
XX fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
XX immune disease caused by an infectious agent. This polynucleotide
XX represents a DNA sequence relating to the genes used in the analysis and
XX treatment of autoimmune diseases or arthritides. Note: This sequence is
XX not shown in the specification. It has been supplied in an electronic
XX format from WIPO.
XX
SQ Sequence 6217 BP; 1647 A; 1404 C; 1702 G; 1464 T; 0 U; 0 Other;

  Query Match      61.3%; Score 19; DB 11; Length 6217;
  Best Local Similarity 81.5%; Pred. No. 4.1e+02;
  Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGGTCTTACGGGTGATGATGTTT 30
```

Db 1546 TTCGGCCTTCGGGTGATCCAGGTTT 1572
|||||
RESULT 28
ADP65717
ID ADP65717 standard; DNA; 6217 BP.
XX AC
XX ADP65717;
XX 12-AUG-2004 (first entry)
XX DT
XX DE Human collagen, type V, alpha 2 (COL5A2) DNA.
XX KW autoimmune disease; arthritis; gene expression analysis;
KW rheumatoid arthritis; collagen-induced; immunosuppressive; antirheumatic;
KW antiarthritic; osteopathic; angiot; antiinflammatory; dermatological;
KW immunomodulatory; lupus; ankylosing spondylitis; Fibrositis;
KW fibromyalgia; osteoarthritis; gout; juvenile rheumatoid arthritis;
KW immune; ds; human.
XX OS Homo sapiens.
XX PN WO2003072827-A1.
XX PD
XX PF 04-SEP-2003.
XX PF 31-OCT-2002; 2002WO-US035433.
XX PF 31-OCT-2001; 2001US-0336220P.
XX PR (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX PA Hirsch R, Thorton SL;
XX PI WPI; 2003-712740/67.
XX DR GENBANK; NW_000393.
XX DR
XX PT Diagnosing and analyzing autoimmune disease using gene expression
PT profiles and microarray technology, useful for diagnosing and treating
PT rheumatoid arthritis, lupus, fibrositis, osteoarthritis, fibromyalgia and
PT gout.
XX PS Disclosure; Page; 56pp; English.
XX PS
XX CC The invention relates to a novel method for diagnosing and analysing
CC autoimmune disease or arthritides. The method comprises obtaining a
CC patient sample containing mRNA, analysing gene expression using the mRNA
CC that results in a gene expression signature of the mRNA, and using that
CC gene expression signature to diagnose or analyse the autoimmune disease
CC or arthritides in the patient, where gene expression of at least 60% of
CC the genes correlates with that of the gene signature. The invention
CC further comprises: a treatment of rheumatoid arthritis; identification of
CC genes for targeting in the treatment of rheumatoid arthritis in a mammal
CC other than a mouse; diagnosis of rheumatoid arthritis in a mammal; an
CC array or gene chip, specific for rheumatoid arthritis; diagnosis or
CC analyses of autoimmune disease or rheumatoid arthritis; screening the
CC efficacy of a candidate drug in vitro for the treatment of collagen-
CC induced arthritis; and reducing the symptoms associated with collagen-
CC activities: immunosuppressive, antirheumatic, antiarthritic, osteopathic,
CC antigout, antiinflammatory, dermatological, and immunomodulatory. The
CC methods and compositions of the present invention are useful for
CC diagnosing and treating autoimmune disease or arthritides, such as
CC rheumatoid arthritis, lupus, ankylosing spondylitis, fibrositis,
CC fibromyalgia, osteoarthritis, gout, juvenile rheumatoid arthritis, and an
CC immune disease caused by an infectious agent. This polynucleotide
CC represents a DNA sequence relating to the genes used in the analysis and
CC treatment of autoimmune diseases or arthritides. Note: This sequence is
CC not shown in the specification. It has been supplied in an electronic
CC format from WIPO.
XX SQ Sequence 6217 BP; 1647 A; 1404 C; 1702 G; 1464 T; 0 U; 0 Other;

Query Match 61.3%; Score 19; DB 11; Length 6217;
Best Local Similarity 81.5%; Pred. No. 4.1e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 TTCGTGCTTACGGGTGATGTTT 30
|||||
Db 1546 TTCGGCCTTCGGGTGATCCAGGTTT 1572
|||||
RESULT 29
ABL29978/c
ID ABL29978 standard; DNA; 6792 BP.
XX AC
XX ABL29978;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 41407.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX PS Claim 1; SEQ ID NO 41407; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 6792 BP; 2149 A; 1373 C; 1288 G; 1982 T; 0 U; 0 Other;
Query Match 61.3%; Score 19; DB 4; Length 6792;
Best Local Similarity 81.5%; Pred. No. 4.1e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 5 TCTGCTTACGGGTGATGTTT 31
|||||
Db 4116 TCTGCTTACGGGTGATGTTT 4090
|||||
RESULT 30
ABV31805
ID ABV31805 standard; cDNA; 211 BP.
XX AC ABV31805;


```
XX 16-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 31796.
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 6820; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing whether prostate cancer has metastasized in a patient; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 211 BP; 38 A; 55 C; 68 G; 50 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 5; Length 211;
Best Local Similarity 76.7%; Pred. No. 3.2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ACGTTCTGGTCTTACGGGTGATGTTT 30
DB 81 ACGTTCTGGGATTCGGGAGATCGGTTT 110
RESULT 31
ABV12481
ID ABV12481 standard; cDNA; 407 BP.
XX
XX ABV12481;
AC
XX 13-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 12472.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
```

```
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 2056; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing whether prostate cancer has metastasized in a patient; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 407 BP; 99 A; 105 C; 104 G; 99 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 5; Length 407;
Best Local Similarity 76.7%; Pred. No. 3.4e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ACGTTCTGGTCTTACGGGTGATGTTT 30
DB 20 ACGTTCTGGGATTCGGGAGATCGGTTT 49
RESULT 32
ABV03312
ID ABV03312 standard; cDNA; 444 BP.
XX
XX ABV03312;
AC
XX 13-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 3303.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
```

PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 603; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a patient
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 444 BP; 103 A; 117 C; 119 G; 104 T; 0 U; 1 Other;
SQ

Query Match 60.6%; Score 18.8; DB 5; Length 444;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTTCGTCTTACGGGTGATGATGTTT 30
||||||| ||||| ||||| |||||
Db 62 ACGTTCGGGATTCGGGAGATCGCGTTT 91

RESULT 33
ABV42537
ID ABV42537 standard; cDNA; 446 BP.
XX
XX ABV42537;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 42528.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PI Schlegel R, Endege WO, Monahan JE;
XX
XX

DR WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 8514; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 446 BP; 106 A; 116 C; 119 G; 105 T; 0 U; 0 Other;
SQ

Query Match 60.6%; Score 18.8; DB 5; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTTCGTCTTACGGGTGATGATGTTT 30
||||||| ||||| ||||| |||||
Db 62 ACGTTCGGGATTCGGGAGATCGCGTTT 91

RESULT 34
ABV33617
ID ABV33617 standard; cDNA; 446 BP.
XX
XX ABV33617;
XX
XX 16-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 33608.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
XX
XX 16-MAR-2000; 2000US-0189862P.
XX
XX 25-MAY-2000; 2000US-0207454P.
XX
XX 09-JUN-2000; 2000US-0211314P.
XX
XX 18-JUL-2000; 2000US-0219007P.
XX
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 7108; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (i) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 446 BP; 106 A; 116 C; 119 G; 105 T; 0 U; 0 Other;

Query Match 60.6%; Score 18.8; DB 5; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACGGGTGATGTAGGTTT 30
Db ||||| ||||| ||||| ||||| ||||| ||||| |||||

62 ACCTTCTGGGATTCGGGAGATCGGGTTT 91

RESULT 35.

ABV42534
ID ABV42534 standard; cDNA; 446 BP.

XX AC ABV42534;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 42525.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 8514; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (i) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (i) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 446 BP; 106 A; 116 C; 119 G; 105 T; 0 U; 0 Other;

Query Match 60.6%; Score 18.8; DB 5; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACGGGTGATGTAGGTTT 30
Db ||||| ||||| ||||| ||||| ||||| ||||| |||||

62 ACCTTCTGGGATTCGGGAGATCGGGTTT 91

RESULT 36

ABV40772

ID ABV40772 standard; cDNA; 446 BP.

XX AC ABV40772;

XX DT 16-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 40763.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX PN WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

XX PR 16-MAR-2000; 2000US-0189862P.

XX PR 25-MAY-2000; 2000US-0207454P.

XX PR 09-JUN-2000; 2000US-0211314P.

XX PR 18-JUL-2000; 2000US-0219007P.

XX PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JE;

XX DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 8213; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (i) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (i) is useful for: (a) assessing whether
XX a patient is afflicted with prostate cancer; (b) monitoring the
XX progression of prostate cancer in a patient; (c) assessing the efficacy
XX of a test compound to inhibit prostate cancer in a patient; (d) assessing
XX the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound; (g)
XX determining whether prostate cancer has metastasized in a patient; (h)
XX assessing the aggressiveness or indolence of prostate cancer in a patient
XX ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 446 BP; 106 A; 116 C; 119 G; 105 T; 0 U; 0 Other;

Query Match 60.6%; Score 18.8; DB 5; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACCGGTGATAGGTTT 30
|||||
Db 62 ACCTTCTGGGATTTCGGGAGATCGGTTT 91

RESULT 37
ACAS2379
ID ACAS2379 standard; DNA; 462 BP.
XX
AC ACAS2379;
XX
DT 19-JUN-2003 (first entry)
XX
DE Prokaryotic essential gene #34036.
XX
KW Antisense; ds; prokaryotic essential gene; cell proliferation;
KW drug design; gene.
XX
OS Treponema pallidum.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Irawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR P-PSDB; ABU48509.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 40249; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway;
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC on a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

CC prokaryotic essential genes. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 462 BP; 91 A; 92 C; 170 G; 109 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 8; Length 462;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 CGTCTCGTCTTACCGGTGATAGGTTT 31
|||||
Db 121 CGTGGCGGCATAAGGCTTATGTAGGTTT 150

RESULT 38
ABN22910/c
ID ABN22910 standard; cDNA; 489 BP.
XX
AC ABN22910;
XX
DT 24-JUN-2002 (first entry)
XX
DE Human ORFX polynucleotide sequence SEQ ID NO:14297.
XX
KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192523-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US010836.
XX
PR 30-MAY-2000; 2000US-0206132P.
PR 29-AUG-2000; 2000US-0228716P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Shinketsu RA, Leach MD;
XX
XX WPI; 2002-106308/14.
XX P-PSDB; ABP07158.
XX
XX Novel human polypeptides and polynucleotides useful for diagnosing,
PT preventing and treating cardiovascular disease, neurodegenerative,
PT hyperproliferative disorders and autoimmune disorders.
XX
XX Disclosure; SEQ ID NO 14297; 1037pp; English.
XX
CC The present invention describes substantially purified human proteins
CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
CC in the specification). ABN15762 to ABN27252 encode the human ORFX
CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
CC treating or preventing a pathology associated with an ORFX-associated
CC disorder in humans, and in the manufacture of a medicament for treating a
CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
CC sequences can be used in gene therapy. ORFX sequences can be used in the
CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, diabetes mellitus, systemic
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 489 BP; 167 A; 109 C; 99 G; 112 T; 0 U; 2 Other;

Query Match 60.6%; Score 18.8; DB 6; Length 489;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 CGTCTCGTCTTACGGGTGATGTAGGTTT 31
DB 83 CGTCTCGTCTTACGGGTGATGTAGGTTT 54

RESULT 39
ABV10635
ID ABV10635 standard; cDNA; 496 BP.
XX
AC ABV10635;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 10626.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX
FN WO200160860-A2.
XX
PD 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JB;
XX
XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 1716; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or incidence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 496 BP; 122 A; 110 C; 106 G; 158 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 5; Length 496;
Best Local Similarity 76.7%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 ACGTCTCGTCTTACGGGTGATGTAGGTTT 30
DB 20 ACGTCTCGGATTCGGGAGATGCGGTTT 49

RESULT 40
ABQ38090
ID ABQ38090 standard; DNA; 751 BP.
XX
AC ABQ38090;
XX
DT 12-JUL-2002 (first entry)
XX

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 24691.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.

XX Homo sapiens.
XX
XX WO200218632-A2.
XX
PD 07-MAR-2002.
XX
XX 01-SEP-2001; 2001WO-EP010074.
XX
XX 01-SEP-2000; 2000DE-01043826.
PR 05-SEP-2000; 2000DE-01044543.
XX
XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for
diagnosis and prognosis, comprises selective hybridization of amplicons
from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP/s); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 751 BP; 104 A; 62 C; 243 G; 337 T; 0 U; 5 Other;

Query Match 60.6%; Score 18.8; DB 6; Length 751;
Best Local Similarity 76.7%; Pred. No. 3.7e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 CGTTCTGCTTACGGGTGATGTAGTGT 31
Db 5 CGTTATAGTTTTTCGGGAGACGTGGGT 34

Search completed: June 4, 2005, 07:29:02
Job time : 237.976 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:48:59 ; Search time 68.9704 Seconds
(without alignments)
735.454 Million cell updates/sec

Title: US-09-674-277-25
Perfect score: 31
Sequence: 1 acgtctgtcttacgggtgatgtagtttt 31

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTCUS COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.4	65.8	1023	4	US-09-543-681A-2804
2	20	64.5	18355	4	US-08-956-171E-67
3	20	64.5	18355	4	US-08-781-986A-67
4	19	61.3	2853	4	US-09-252-991A-12441
5	19	61.3	3864	4	US-09-252-991A-13035
6	18.8	60.6	3811	4	US-09-112-580-5
7	18.8	60.6	5076	4	US-09-949-016-1616
8	18.6	60.0	30502	4	US-09-949-016-15873
9	18.4	59.4	171	3	US-09-080-983-6
10	18.4	59.4	1868	4	US-08-956-171E-101
11	18.4	59.4	1868	4	US-08-781-986A-101
12	18.4	59.4	1869	4	US-09-148-545-56
13	18.4	59.4	1960	4	US-09-148-545-114
14	18.4	59.4	3078	4	US-09-543-681A-2940
15	18.4	59.4	15500	3	US-09-080-983-1
16	18.4	59.4	150394	4	US-09-949-016-13042
17	18.2	58.7	699	4	US-09-328-352-3899
18	18.2	58.7	35609	4	US-09-949-016-17370
19	18	58.1	396	4	US-09-248-796A-1391
20	18	58.1	601	4	US-09-949-016-201955
21	18	58.1	133719	4	US-09-949-016-15092
22	18	58.1	137753	4	US-09-949-016-17404
23	18	58.1	678533	4	US-09-949-016-14577
24	18	58.1	678533	4	US-09-949-016-14578
25	17.8	57.4	601	4	US-09-949-016-144937
26	17.8	57.4	728	3	US-09-221-017B-677
27	17.8	57.4	729	3	US-08-858-207A-198

28	17.8	57.4	879	4	US-09-583-110-132	Sequence 132, App
29	17.8	57.4	888	4	US-09-107-433-2440	Sequence 2440, Ap
30	17.8	57.4	2193	4	US-09-328-352-3985	Sequence 3985, Ap
31	17.8	57.4	5155	4	US-09-902-540-755	Sequence 755, Appl
32	17.8	57.4	8654	3	US-08-961-527-98	Sequence 98, Appl
33	17.8	57.4	85122	4	US-09-949-016-14693	Sequence 14693, A
34	17.8	57.4	119214	4	US-09-949-016-12507	Sequence 12507, A
35	17.8	57.4	177251	4	US-09-949-016-15841	Sequence 15841, A
36	17.8	57.4	236474	4	US-09-949-016-13418	Sequence 13418, A
37	17.4	56.1	552	4	US-09-248-796A-13671	Sequence 13671, A
38	17.4	56.1	597	4	US-09-328-352-1942	Sequence 1942, Ap
39	17.4	56.1	601	4	US-09-949-016-152745	Sequence 152745, A
40	17.4	56.1	915	4	US-09-640-211A-55	Sequence 55, Appl
41	17.4	56.1	1107	4	US-09-902-540-6753	Sequence 6753, Ap
42	17.4	56.1	1110	4	US-09-134-000C-2300	Sequence 2300, Ap
43	17.4	56.1	2780	4	US-09-841-786-10	Sequence 10, Appl
44	17.4	56.1	3707	4	US-09-902-540-549	Sequence 549, App
45	17.4	56.1	9726	4	US-09-841-786-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-543-681A-2804
; Sequence 2804, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 2804
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-2804

Query Match 65.8%; Score 20.4; DB 4; Length 1023;
Best Local Similarity 80.0%; Pred. No. 15;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	ACGTTCTGCTTACGGGTGATGTTT	30
DB	183	ACGTTCTGCTTACGGGTGATGTTT	212

RESULT 2

US-08-956-171E-67
; Sequence 67, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSER: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII Text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/956,171E
;; FILING DATE: 20-Oct-1997
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/009,861
;; FILING DATE: January 5, 1996
;; APPLICATION NUMBER: 08/781,986
;; FILING DATE: January 3, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mark J. Hyman
;; REGISTRATION NUMBER: 46,789
;; REFERENCE/DOCKET NUMBER: PB248P1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (240) 314-1224
;; TELEFAX: (301) 309-8439
;; INFORMATION FOR SEQ ID NO: 67:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 18355 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; SEQUENCE DESCRIPTION: SEQ ID NO: 67:

Query Match 64.5%; Score 20; DB 4; Length 18355;
Best Local Similarity 82.1%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGAGTGGTTT 31
||||||| ||||| ||||| ||||| |||||
Db 14008 TTCTGTCATATGATGAGTGGTTT 14035

RESULT 3
US-08-781-986A-67
; Sequence 67, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 67:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 18355 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; US-08-781-986A-67

Query Match 64.5%; Score 20; DB 4; Length 18355;
Best Local Similarity 82.1%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGAGTGGTTT 31
||||||| ||||| ||||| ||||| |||||
Db 14008 TTCTGTCATATGATGAGTGGTTT 14035

RESULT 4
US-09-252-991A-12441
; Sequence 12441, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12441
; LENGTH: 2853
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-12441

Query Match 61.3%; Score 19; DB 4; Length 2853;
Best Local Similarity 81.5%; Pred. No. 74;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CGTCTGCTCTTACGGGTGATGAGT 28
||| ||||| ||||| ||||| |||||
Db 910 CGCGCGGCTCTTACCGGTGATGTTGTT 936

RESULT 5
US-09-252-991A-13035/c
; Sequence 13035, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13035
; LENGTH: 3864
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-13035

Query Match 61.3%; Score 19; DB 4; Length 3864;
Best Local Similarity 81.5%; Pred. No. 78;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;


```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-080-983-6

Query Match          59.4%; Score 18.4; DB 3; Length 171;
Best Local Similarity 78.6%; Pred. No. 82;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGATGTTT 31
    |||||
Db 77 TTCTGCTCTTCCGGGTGATTAAGTCTTT 104
    |||||

RESULT 10
US-08-956-171E-101/c
; Sequence 101, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 101:
US-08-956-171E-101

Query Match          59.4%; Score 18.4; DB 4; Length 1868;
Best Local Similarity 78.6%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGATGTTT 31
    |||||
Db 1267 TGCTGTATTACTGGGTGTGAAGTTT 1240
    |||||

RESULT 11
US-08-781-986A-101/c
; Sequence 101, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1868 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-781-986A-101

Query Match          59.4%; Score 18.4; DB 4; Length 1868;
Best Local Similarity 78.6%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGATGTTT 31
    |||||
Db 1267 TGCTGTATTACTGGGTGTGAAGTTT 1240
    |||||

RESULT 12
US-09-148-545-56/c
; Sequence 56, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; CURRENT FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
```


EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 56
LENGTH: 1869

Query Match 59.4%; Score 18.4; DB 4; Length 1869;

Best Local Similarity 78.6%; Pred. No. 1.3e+02;

Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 TTCTGGTCTTACGGTGATGATAGGTTTTT 31
Db 757 TTTCGGTATCTCGGTGATGATAGGTTAT 730

RESULT 13

US-09-148-545-114/c
Sequence 114, Application US/09148545
Patent No 6590075
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 70 Human Secreted Proteins
FILE REFERENCE: P2001P1
CURRENT APPLICATION NUMBER: US/09/148,545
CURRENT FILING DATE: 1998-09-04
EARLIER APPLICATION NUMBER: PCT/US98/04482
EARLIER FILING DATE: 1998-03-06
EARLIER APPLICATION NUMBER: 60/040,162
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,333
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/038,621
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,161
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,626
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,334
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,336
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/040,163
EARLIER FILING DATE: 1997-03-07
EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,600
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,597
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,502
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,633
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,583
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,617
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,618
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,503
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,592
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,581
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,584
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,500
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,587
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,492
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,598
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,613
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,582
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,596
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,612
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,632
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,601
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,580
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,568
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,314
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,569
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,311
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,671
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,674
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,669
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,312
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,313
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,672
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,315
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/056,886
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,877
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,889
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,893
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,630
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,878
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,662
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,872
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,882
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,637
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,903
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,888
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,879
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,880
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,894
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,911
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,636
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,874
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,910
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,864
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,631
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,845
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/057,761
EARLIER FILING DATE: 05-Sep-1997
EARLIER APPLICATION NUMBER: 60/047,599
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,588
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,585
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,586
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,590
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,594
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,589
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,593
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/047,614
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,578
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/043,576
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/047,501
EARLIER FILING DATE: 1997-05-23
EARLIER APPLICATION NUMBER: 60/043,670
EARLIER FILING DATE: 1997-04-11
EARLIER APPLICATION NUMBER: 60/056,632

EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,664
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,876
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,881
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,909
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,875
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,862
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,887
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/056,908
EARLIER FILING DATE: 1997-08-22
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/057,650
EARLIER FILING DATE: 1997-09-05
EARLIER APPLICATION NUMBER: 60/056,884
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 280
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 114
LENGTH: 1960
Query Match 59.4%; Score 18.4; DB 4; Length 1960;
Best Local Similarity 78.6%; Pred. No. 1.3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 TTCTGGTCTTACGGGTGATGTAGGTTT 31
DB 798 TTTCGTATCTCGGTGATGTAGTTAT 771
RESULT 14
US-09-543-681A-2940/c
Sequence 2940, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILI
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 2940
LENGTH: 3078
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2940
Query Match 59.4%; Score 18.4; DB 4; Length 3078;
Best Local Similarity 78.6%; Pred. No. 1.4e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 TTCTGGTCTTACGGGTGATGTAGGTTT 31
DB 768 TTTCGTGTTTACAGGTAATTGAGTTGT 741
RESULT 15
US-09-080-983-1
Sequence 1, Application US/09080983
Patent No. 6197948
GENERAL INFORMATION:
APPLICANT: Zhu, Hai-Ying
APPLICANT: Ling, Kai-Shu

APPLICANT: Gonsalves, Dennis
TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS TYPE 2 PROTEINS
AND THEIR USES
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/080,983
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,194
FILING DATE: 20-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1631
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 15500 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-080-983-1

Query Match 59.4%; Score 18.4; DB 3; Length 15500;
Best Local Similarity 78.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 TTCTGCTCTTACGGGTGATGTAGGTTTT 31
|||||
Db 9441 TTCTGCTCTCCGGGTGATTAAGTCTTT 9468

RESULT 16
US-09-949-016-13042/c
Sequence 13042, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13042
LENGTH: 150394
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature

LOCATION: (1)...(150394)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13042
Query Match 59.4%; Score 18.4; DB 4; Length 150394;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 TTCTGCTCTTACGGGTGATGTAGGTTTT 31
|||||
Db 111458 TTCTGCTCTTACGGGTGATGAAGTGTT 111431
RESULT 17
US-09-328-352-3899
Sequence 3899, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 3899
LENGTH: 699
TYPE: DNA
ORGANISM: Acinetobacter baumannii
US-09-328-352-3899
Query Match 58.7%; Score 18.2; DB 4; Length 699;
Best Local Similarity 87.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 GTTCGTCTTACGGGTGATGTA 25
|||||
Db 659 GTCTGCTCTTACTGTTGATGTA 681

RESULT 18
US-09-949-016-17370/c
Sequence 17370, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17370
LENGTH: 35609
TYPE: DNA
ORGANISM: Human
US-09-949-016-17370

Query Match 58.7%; Score 18.2; DB 4; Length 35609;
Best Local Similarity 74.2%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 ACGTTCTGCTCTTACGGGTGATGTAGGTTTT 31
|||||
Db 21821 ACGTTACGTTTACTGTTATTTGGGTTTT 21791

```

RESULT 19
US-09-248-796A-1391
; Sequence 1391, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 1391
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-1391

Query Match      58.1%; Score 18; DB 4; Length 396;
Best Local Similarity 80.8%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TCTGGTCTTACGGGTGATGTAGTTT 30
Db 118 TATGCTATTACTGCTGATGTAGTTT 143

RESULT 20
US-09-949-016-201955
; Sequence 201955, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 201955
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201955

Query Match      58.1%; Score 18; DB 4; Length 601;
Best Local Similarity 80.8%; Pred. No. 1.5e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TCTGGTCTTACGGGTGATGTAGTTT 30
Db 268 TTGGTCTTTAGGCTGATGTAGGCTT 293

RESULT 21
US-09-949-016-15092/c
; Sequence 15092, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 15092
; LENGTH: 133719
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15092

Query Match      58.1%; Score 18; DB 4; Length 133719;
Best Local Similarity 80.8%; Pred. No. 4.1e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TCTGGTCTTACGGGTGATGTAGTTT 30
Db 94987 TTGGTCTTTAGGCTGATGTAGGCTT 95012

RESULT 22
US-09-949-016-17404
; Sequence 17404, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 17404
; LENGTH: 137753
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17404

Query Match      58.1%; Score 18; DB 4; Length 137753;
Best Local Similarity 80.8%; Pred. No. 4.1e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TCTGGTCTTACGGGTGATGTAGTTT 30
Db 94987 TTGGTCTTTAGGCTGATGTAGGCTT 95012

RESULT 23
US-09-949-016-14577/c
; Sequence 14577, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 14577
; LENGTH: 133719
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14577/c

Query Match      58.1%; Score 18; DB 4; Length 133719;
Best Local Similarity 80.8%; Pred. No. 4.1e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TCTGGTCTTACGGGTGATGTAGTTT 30
Db 94987 TTGGTCTTTAGGCTGATGTAGGCTT 95012

```

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14577
; LENGTH: 678533
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(678533)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14577
```

```
Query Match 58.1%; Score 18; DB 4; Length 678533;
Best Local Similarity 80.8%; Pred. No. 4.9e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 2 CTTCTGCTCTTACGGGTGATGTAGG 27
| | | | | | | | | | | | | | | | | |
Db 648690 CTTTCTGTTCTTAGGGGTGTGAAGG 648665
```

```
RESULT 24
US-09-949-016-14578/c
; Sequence 14578, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14578
; LENGTH: 678533
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(678533)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14578
```

```
Query Match 58.1%; Score 18; DB 4; Length 678533;
Best Local Similarity 80.8%; Pred. No. 4.9e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 2 CTTCTGCTCTTACGGGTGATGTAGG 27
| | | | | | | | | | | | | | | | | |
Db 648690 CTTTCTGTTCTTAGGGGTGTGAAGG 648665
```

```
RESULT 25
US-09-949-016-144937
; Sequence 144937, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144937
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-144937
```

```
Query Match 57.4%; Score 17.8; DB 4; Length 601;
Best Local Similarity 75.9%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 2 CTTCTGCTCTTACGGGTGATGTAGGTTT 30
| | | | | | | | | | | | | | | | | |
Db 8 CTTCTGCTCTTAATTGGTCTGTGGGTTT 36
```

```
RESULT 26
US-09-221-017B-677
; Sequence 677, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
```


; INFORMATION FOR SEQ ID NO: 677:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 728 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...728
; US-09-221-017B-677

Query Match . 57.4%; Score 17.8; DB 3; Length 728;
Best Local Similarity 75.9%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGTCTGCTTACGGGTGATGAGTGT 30
||| ||||| ||||| ||||| |||||
DB 63 CGTAATGCTCTGTGATGATGACGTT 91

RESULT 27
US-08-858-207A-198
; Sequence 198, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 198:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-858-207A-198

Query Match 57.4%; Score 17.8; DB 3; Length 729;

Best Local Similarity 75.9%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 GTTCTGGTCTTACGGGTGATGAGTGT 31
||||| ||||| ||||| ||||| |||||
DB 385 GTTCTGCTCTTGGCGGTGCGGCTGT 413

RESULT 28
US-09-583-110-132
; Sequence 132, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: Pneumoniae for Diagnostics and Therapeutics
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 132
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; US-09-583-110-132

Query Match 57.4%; Score 17.8; DB 4; Length 879;
Best Local Similarity 75.9%; Pred. No. 2e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GTTCTGCTTACGGGTGATGAGTGT 31
||||| ||||| ||||| ||||| |||||
DB 587 GTTCTGCTCTTGGCGGTGCGGCTGT 615

RESULT 29
US-09-107-433-2440
; Sequence 2440, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinllo, Pamela Deneke

```
/
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-011
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 2440:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 888 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Streptococcus pneumoniae
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (B) LOCATION 1...888
/ SEQUENCE DESCRIPTION: SEQ ID NO: 2440:
US-09-107-433-2440

Query Match          57.4%; Score 17.8; DB 4; Length 888;
Best Local Similarity 75.9%; Pred. No. 2e+02; 7; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GTTCTGGTCTTACGGGTGATGAGTTT 31
   ||||| ||||| ||||| ||||| |||||
Db 596 GTTCTGCTCTTGGGGTGGCGTGGT 624

RESULT 30
US-09-328-352-3985
/ Sequence 3985, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 3985
/ LENGTH: 2193
/ TYPE: DNA
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-3985

Query Match          57.4%; Score 17.8; DB 4; Length 2193;
Best Local Similarity 75.9%; Pred. No. 2.4e+02; 7; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGTCTGGTCTTACGGGTGATGAGTTT 30
   ||||| ||||| ||||| ||||| |||||
Db 1780 CGTCTGGTCAAAGAATGATGAGTTT 1808

RESULT 31
US-09-902-540-755
/ Sequence 755, Application US/09902540
/ Patent No. 6833447
/ GENERAL INFORMATION:
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Wiegand, Roger C.
/ TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
/ FILE REFERENCE: 38-10(15849)B
/ CURRENT APPLICATION NUMBER: US/09/902,540
/ CURRENT FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/217,883
/ PRIOR FILING DATE: 2000-07-10
```

```
/
/ NUMBER OF SEQ ID NOS: 16825
/ SEQ ID NO 755
/ LENGTH: 5155
/ TYPE: DNA
/ ORGANISM: Myxococcus xanthus
US-09-902-540-755

Query Match          57.4%; Score 17.8; DB 4; Length 5155;
Best Local Similarity 75.9%; Pred. No. 2.8e+02; 7; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GTTCTGGTCTTACGGGTGATGAGTTT 31
   ||||| ||||| ||||| ||||| |||||
Db 4184 GTTCTGCTCTCCGGTGGGGTGGTCTT 4212

RESULT 32
US-08-961-527-98
/ Sequence 98, Application US/08961527
/ Patent No. 6420135
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 391
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/961,527
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brookes, A. Anders
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: PB340P1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 98:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 8654 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
US-08-961-527-98

Query Match          57.4%; Score 17.8; DB 3; Length 8654;
Best Local Similarity 75.9%; Pred. No. 3.1e+02; 7; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GTTCTGGTCTTACGGGTGATGAGTTT 31
   ||||| ||||| ||||| ||||| |||||
Db 2016 GTTCTGCTCTTGGGGTGGCGTGGT 2044

RESULT 33
US-09-949-016-14693/c
/ Sequence 14693, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
```

```
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14693
; LENGTH: 85122
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(85122)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14693

Query Match          57.4%; Score 17.8; DB 4; Length 85122;
Best Local Similarity 75.9%; Pred. No. 4.6e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GTTCTGCTCTTACGGGTGATGTAGTGT 31
   ||||| ||||| ||||| ||||| |||||
Db 26357 GTTCTGTTTACAGTTTATCTAGTTT 26329

RESULT 34
US-09-949-016-12507/c
; Sequence 12507, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12507
; LENGTH: 119214
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(119214)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12507

Query Match          57.4%; Score 17.8; DB 4; Length 119214;
Best Local Similarity 75.9%; Pred. No. 4.9e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GTTCTGCTCTTACGGGTGATGTAGTGT 31
   ||||| ||||| ||||| ||||| |||||
Db 61000 GTTCTGTTTACAGTTTATCTAGTTT 60972

RESULT 35
US-09-949-016-15841
; Sequence 15841, Application US/09949016
```

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15841
; LENGTH: 177251
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-15841

Query Match          57.4%; Score 17.8; DB 4; Length 177251;
Best Local Similarity 75.9%; Pred. No. 5.2e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CTTCTGCTCTTACGGGTGATGTAGTGT 30
   ||||| ||||| ||||| ||||| |||||
Db 77577 CTTCTGCTCTTAAATGGTCTGTGGT 77605

RESULT 36
US-09-949-016-13418
; Sequence 13418, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13418
; LENGTH: 236474
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc_feature
; LOCATION: (1)...(236474)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13418

Query Match          57.4%; Score 17.8; DB 4; Length 236474;
Best Local Similarity 75.9%; Pred. No. 5.4e+02;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 GTTCTGCTCTTACGGGTGATGTAGTGT 31
   ||||| ||||| ||||| ||||| |||||
Db 27793 GGTGTGGTCTTGTGGGTGATGAGGAT 27821

RESULT 37
US-09-248-796A-13671/c
; Sequence 13671, Application US/09248796A
; Patent No. 6747137
```

```
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 13671
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-13671

Query Match          56.1%; Score 17.4; DB 4; Length 552;
Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TCTGCTCTTACGGGTGATGTAGTGT 31
Db 428 TCTGGTGTCTGGTGTGATGTAGGCATT 402

RESULT 38
US-09-328-352-1942
; Sequence 1942, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1942
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1942

Query Match          56.1%; Score 17.4; DB 4; Length 597;
Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CGTCTCTGCTTACGGGTGATGTAGGT 28
Db 139 CGTCTCTGCTTAAATGGGTGTGTGCT 165

RESULT 39
US-09-949-016-152745
; Sequence 152745, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 152745
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-152745

Query Match          56.1%; Score 17.4; DB 4; Length 601;
Best Local Similarity 77.8%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 TCTGCTCTTACGGGTGATGTAGTGT 31
Db 501 TCTGGTCTTTGGGTGATTTGTATTT 527

RESULT 40
US-09-640-211A-55
; Sequence 55, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Sherk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 915
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-09-640-211A-55

Query Match          56.1%; Score 17.4; DB 4; Length 915;
Best Local Similarity 77.8%; Pred. No. 3.1e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AGTTTCTGCTTACGGGTGATGTAGG 27
Db 551 AAGTCTGTCATATGCGGTGTGTGG 577

Search completed: June 4, 2005, 11:53:15
Job time : 72.9704 secs
```


; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7056
; LENGTH: 23865
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-7056

Query Match 65.8%; Score 20.4; DB 18; Length 23865;
Best Local Similarity 80.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CGTTCGCTTACGGGTGATGTAGGTTTT 31
Db 17757 CTTTGATGCTTAGGGGAGATGTAGGTTTT 17728

RESULT 3
US-10-719-993-6873
; Sequence 6873, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6873
; LENGTH: 65273
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-719-993-6873

Query Match 65.8%; Score 20.4; DB 18; Length 65273;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 CGTTCGCTTACGGGTGATGTAGGTTTT 31
Db 60524 CTTTGATGCTTAGGGGAGATGTAGGTTTT 60553

RESULT 4
US-10-425-115-43831/c
; Sequence 43831, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 43831
; LENGTH: 373
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_13997C.1
US-10-425-115-43831

Query Match 65.2%; Score 20.2; DB 18; Length 373;
Best Local Similarity 88.0%; Pred. No. 80;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 CTGGTCTTACGGGTGATGTAGGTTTT 30
Db 206 CTGGTCTTACGGGTGATGTAGGATT 182

RESULT 5
US-10-027-632-16497
; Sequence 16497, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16497
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-16497

Query Match 64.5%; Score 20; DB 13; Length 740;
Best Local Similarity 82.1%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGGTCTTACGGGTGATGTAGGTTTT 31
Db 489 TTTTGTGTTTTACGGGTGATATTGGTTTT 516

RESULT 6
US-10-027-632-16497
; Sequence 16497, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```

; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16497
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-16497

```

Query Match	64.5%	Score 20;	DB 17;	Length 740;
Best Local Similarity	82.1%	Pred. No. 1.1e-02;		
Matches 23;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
4	TTCTGCTCTTACGGGTGATGTAGGTTTT	31		
489	TTTTTTTTTACGGGTGATGTAGGTTTT	516		

```

RESULT 7
US-08-781-986A-67
; Sequence 67, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A

```

Query Match	64.5%	Score 20	DB 8	Length 18355
Best Local Similarity	82.1%	Pred. No. 1.7e+02		
Matches: 23	Conservative 0	Mismatches 5	Indels 0	Gaps 0
Qy	4	TTCTGGTCTTACGGGTGATGAGTTTT	31	
Db	14008	TTCTGGTCAATGATGAAGTTGTTTT	14035	

RESULT 8
US-10-329-624-67

```

; Sequence 67, Application US/10329624
; Publication No. US20040043037A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
;           Gil H. Choi
;           Patrick S. Dillon
;           Craig A. Rosen
;           Steven C. Barash
;           Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/329,624
; FILING DATE: 27-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/956,171
; FILING DATE: October 20, 1997
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248PLD1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18355 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 67:
; JS-10-329-624-67

```

Query Match 64.5%; Score 20; DB 17; Length 18355;
Best Local Similarity 82.1%; Pred. No. 1.7e+02;
Matches 23: Conservative 0; Mismatches 5; Indels 0; Gaps 0;

RESULT 9
 US-10-424-599-131231/c
 ; Sequence 131231, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao.Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684

3


```
Db 7360 ACCTTCGGTATTACGTTTGTAGTTT 7330
|||||
RESULT 13
US-10-322-281-706
; Sequence 706, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 706
; LENGTH: 107543
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-706

Query Match 63.9%; Score 19.8; DB 18; Length 107543;
Best Local Similarity 77.4%; Pred. No. 2.6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACCTTCGGTCTTACCGGTGATGATGTTT 31
|||||
Db 65009 ATGGTTGTTGTTACCGGTGATGTTGTTCT 65039
|||||

RESULT 14
US-10-312-841-1
; Sequence 1, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 1
; LENGTH: 367378
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (3294164)
US-10-312-841-1

Query Match 63.9%; Score 19.8; DB 16; Length 367378;
Best Local Similarity 77.4%; Pred. No. 3.8e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACCTTCGGTCTTACCGGTGATGATGTTT 31
|||||
Db 2332382 ACCTTCGGTTTAGGGTTTGTAGTTT 2332412
|||||

RESULT 15
US-10-282-122A-38565
; Sequence 38565, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl

APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 38565
LENGTH: 1581
TYPE: DNA
ORGANISM: Streptococcus pyogenes
US-10-282-122A-38565

Query Match 63.2%; Score 19.6; DB 17; Length 1581;
Best Local Similarity 84.6%; Pred. No. 1.8e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 TTCTGCTTACGGGTGATGATGTTT 29
|||||
Db 243 TTCTGTTGTTACTGCTGATTTAGTTT 268
|||||

RESULT 16
US-10-357-930-1457
; Sequence 1457, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF HUMAN PROSTATE CANCER
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
```

```
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1457
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 46, 81, 159, 160, 162, 188, 198, 222, 230, 231, 238, 244,
; LOCATION: 246, 250, 261, 271, 272, 296, 303, 331, 340, 345, 352, 353,
; LOCATION: 356, 359, 361, 362, 381, 382, 383, 386, 390, 397, 410, 415,
; LOCATION: 454, 456, 462, 463, 468, 482, 503, 510, 517, 524, 535
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 540, 548, 559, 569
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-1457
```

```
Query Match 62.6%; Score 19.4; DB 18; Length 576;
Best Local Similarity 76.7%; Pred. No. 1.9e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 1 ACGTCTCGTCTTACGGGTGATGTAGTTT 30
||||| ||||| ||||| ||||| |||||
DB 20 ACGTCTGGGATTCGGGAGATCGGNTTT 49
```

```
RESULT 17
US-10-437-963-98551/c
; Sequence 98551, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 98551
; LENGTH: 1704
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1704)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MKT4530_96448C.1
US-10-437-963-98551
```

```
Query Match 62.6%; Score 19.4; DB 18; Length 1704;
Best Local Similarity 79.3%; Pred. No. 2.2e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 ACGTCTCGTCTTACGGGTGATGTAGTTT 29
||||| ||||| ||||| ||||| |||||
DB 112 ACATTATGTCCTACATGTGATGTAGTTT 84
```

```
RESULT 18
```

```
US-10-240-485-135
; Sequence 135, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCE: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 135
; LENGTH: 6129
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
```

```
US-10-240-485-135
Query Match 62.6%; Score 19.4; DB 15; Length 6129;
Best Local Similarity 79.3%; Pred. No. 2.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 3 GTTCTGCTCTTACGGGTGATGTAGTTT 31
||||| ||||| ||||| ||||| |||||
DB 1181 GTTGTGTGTTTCGGGGGAGGTAGTTT 1209
```

```
RESULT 19
US-10-282-122A-18012/c
; Sequence 18012, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

```

; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: WRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 4629
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-177-293-69

Query Match 61.3%; Score 19; DB 15; Length 4629;
Best Local Similarity 81.5%; Pred. No. 3.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTTACGGGTGATGAGTTT 30
Db 1527 TTCCGGCCTTCCGGGTGATCCAGGTTT 1553

RESULT 22
US-10-293-582-22
; Sequence 22, Application US/10293582
; Publication No. US20030175253A1
; GENERAL INFORMATION:
; APPLICANT: Akil, Huda
; APPLICANT: Bunney, William E.
; APPLICANT: Burke, Sharon
; APPLICANT: Choudary, Prabhakara V.
; APPLICANT: Cox, David R.
; APPLICANT: Evans, Simon
; APPLICANT: Jones, Edward G.
; APPLICANT: Li, Jun
; APPLICANT: Lopez, Juan P.
; APPLICANT: The Trustees of The Leland Stanford Junior University
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating
; TITLE OF INVENTION: Mental Disorders
; FILE REFERENCE: 020885-000210US
; CURRENT APPLICATION NUMBER: US/10/293,582
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/339,252
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 4629
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: procollagen alpha 2(V), COL5A2, procollagen type V
; US-10-293-582-22

```

```
Query Match      61.3%; Score 19; DB 16; Length 4629;
Best Local Similarity 81.5%; Pred. No. 3.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 TTCTGGTCTTACGGGTGATGTTAGGTTT 30
Db      1527 TTTCGGCCTTCCGGGTGATCCAGGTTT 1553

RESULT 23
US-10-236-031B-73
; Sequence 73, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Millenium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 4629
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-236-031B-73

Query Match      61.3%; Score 19; DB 17; Length 4629;
Best Local Similarity 81.5%; Pred. No. 3.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 TTCTGGTCTTACGGGTGATGTTAGGTTT 30
Db      1527 TTTCGGCCTTCCGGGTGATCCAGGTTT 1553

RESULT 24
US-10-788-792-121
; Sequence 121, Application US/10788792
; Publication No. US20040191819A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Pharmaceuticals Corporation
; APPLICANT: Eveleigh, Deepa
; APPLICANT: Bigwood, Douglas
; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE
; FILE REFERENCE: 5152
; CURRENT APPLICATION NUMBER: US/10/788,792
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: US 60/450,655
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 121
; LENGTH: 4629
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-788-792-121

Query Match      61.3%; Score 19; DB 18; Length 4629;
Best Local Similarity 81.5%; Pred. No. 3.8e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 TTCTGGTCTTACGGGTGATGTTAGGTTT 30
Db      1527 TTTCGGCCTTCCGGGTGATCCAGGTTT 1553

RESULT 25
US-10-301-822-34
; Sequence 34, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF COLON CANCER
; FILE REFERENCE: MPM01-029P2RNM
; CURRENT APPLICATION NUMBER: US/10/301,822
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 6217
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (158)...(4648)
US-10-301-822-34

Query Match      61.3%; Score 19; DB 15; Length 6217;
Best Local Similarity 81.5%; Pred. No. 3.9e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 TTCTGGTCTTACGGGTGATGTTAGGTTT 30
Db      1546 TTTCGGCCTTCCGGGTGATCCAGGTTT 1572

RESULT 26
US-10-027-632-53712/c
; Sequence 53712, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53712
; LENGTH: 715517
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(715517)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-53712

Query Match          61.3%; Score 19; DB 13; Length 715517;
Best Local Similarity 75.9%; Pred. No. 7.3e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACGTTCGTGCTTACCGGTGATGAGTT 29
   ||||||| ||| ||| ||| ||| |||
Db 654932 AYGTTCTGGTTTATGAATGATGGCTT 654904

RESULT 27
US-10-027-632-53712/c
; Sequence 53712, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53712
; LENGTH: 715517
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(715517)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-53712

Query Match          61.3%; Score 19; DB 17; Length 715517;
Best Local Similarity 75.9%; Pred. No. 7.3e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACGTTCGTGCTTACCGGTGATGAGTT 29
   ||||||| ||| ||| ||| ||| |||
Db 654932 AYGTTCTGGTTTATGAATGATGGCTT 654904

RESULT 28
US-10-357-930-31823
; Sequence 31823, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12472
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-12472

; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31823
; LENGTH: 211
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-31823

Query Match          60.6%; Score 18.8; DB 18; Length 211;
Best Local Similarity 76.7%; Pred. No. 3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTTCGTGCTTACCGGTGATGAGTTT 30
   ||||||| ||| ||| ||| ||| |||
Db 81 ACGTTCGTGATTCGGGAGATCGGTTT 110

RESULT 29
US-10-357-930-12472
; Sequence 12472, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12472
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-12472
```

```
Query Match          60.6%; Score 18.8; DB 18; Length 407;
Best Local Similarity 76.7%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTCTCGTCTTACGGGTGATGAGTTT 30
    ||||| ||||| ||||| ||||| |||||
Db 20 ACGTCTCGGATTTCGGGAGATCGCGTTT 49

RESULT 30
US-10-357-930-3303
; Sequence 3303, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: HUMAN PROSTATE CANCER
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3303
; LENGTH: 444
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 211
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-3303

Query Match          60.6%; Score 18.8; DB 18; Length 444;
Best Local Similarity 76.7%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTCTCGTCTTACGGGTGATGAGTTT 30
    ||||| ||||| ||||| ||||| |||||
Db 62 ACGTCTCGGATTTCGGGAGATCGCGTTT 91

RESULT 31
US-10-357-930-33635
; Sequence 33635, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: HUMAN PROSTATE CANCER
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
```

```
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33635
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-33635

Query Match          60.6%; Score 18.8; DB 18; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTCTCGTCTTACGGGTGATGAGTTT 30
    ||||| ||||| ||||| ||||| |||||
Db 62 ACGTCTCGGATTTCGGGAGATCGCGTTT 91

RESULT 32
US-10-357-930-40791
; Sequence 40791, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: HUMAN PROSTATE CANCER
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40791
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-40791

Query Match          60.6%; Score 18.8; DB 18; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTCTCGTCTTACGGGTGATGAGTTT 30
    ||||| ||||| ||||| ||||| |||||
```

Db 62 ACCTTCTGGGATTCGGGAGATCGCGTTT 91

RESULT 33

US-10-357-930-42553
; Sequence 42553, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42553
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-42553

Query Match 60.6%; Score 18.8; DB 18; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACGGGTGATGTAGTTT 30

Db 62 ACCTTCTGGGATTCGGGAGATCGCGTTT 91

RESULT 34

US-10-357-930-42556
; Sequence 42556, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007

; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42556
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-42556

Query Match 60.6%; Score 18.8; DB 18; Length 446;
Best Local Similarity 76.7%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACGGGTGATGTAGTTT 30

Db 62 ACCTTCTGGGATTCGGGAGATCGCGTTT 91

RESULT 35

US-10-282-122A-40249
; Sequence 40249, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40249
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Treponema pallidum
US-10-282-122A-40249

Query Match 60.6%; Score 18.8; DB 17; Length 462;
Best Local Similarity 76.7%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGTTCGTCTTACGGGTGATGATGTTT 31
|||||
Db 121 CGTGGGGCATAAGGCTTATGTAGTGT 150

RESULT 36

US-10-357-930-10626
; Sequence 10626, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10626
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-10626

Query Match 60.6%; Score 18.8; DB 18; Length 496;
Best Local Similarity 76.7%; Pred. No. 3.4e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTTCGTCTTACGGGTGATGATGTTT 30
|||||
Db 20 ACGTTCGGGATTCGGGAGATCGGTTT 49

RESULT 37

US-10-363-345A-24681
; Sequence 24681, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 24681
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 24681
; NAME/KEY: unsure
; LOCATION: (440, 451, 467, 665, 701)
; FEATURE:
US-10-363-345A-24681

; NAME/KEY: unsure
; LOCATION: (440, 451, 467, 665, 701)
US-10-363-345A-24681

Query Match 60.6%; Score 18.8; DB 18; Length 751;
Best Local Similarity 76.7%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGTTCGTCTTACGGGTGATGATGTTT 31
|||||
Db 5 CGTTATAGTTTTTCGGGAGACGTGGTTT 34

RESULT 38

US-10-363-345A-24682/c
; Sequence 24682, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; TITLE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 24682
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 24682
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (51, 87, 285, 301, 312)
US-10-363-345A-24682

Query Match 60.6%; Score 18.8; DB 18; Length 751;
Best Local Similarity 76.7%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGTTCGTCTTACGGGTGATGATGTTT 31
|||||
Db 747 CGTTATAGTTTTTCGGGAGACGTGGTTT 718

RESULT 39

US-10-363-483A-24681
; Sequence 24681, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 24681
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 24681
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (440, 451, 467, 665, 701)
US-10-363-483A-24681

Query Match 60.6%; Score 18.8; DB 19; Length 751;
Best Local Similarity 76.7%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGTTCTGGTCTTACGGGTGATGTAGGTTTT 31
Db 5 CGTTATAGTTTTTCGGGACACGTGGGTTTT 34

RESULT 40
US-10-363-483A-24682/c
; Sequence 24682, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; TITLE OF INVENTION: illnesses
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 24682
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-island No: 24682
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (51, 87, 285, 301, 312)
US-10-363-483A-24682

Query Match 60.6%; Score 18.8; DB 19; Length 751;
Best Local Similarity 76.7%; Pred. No. 3.6e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 CGTTCTGGTCTTACGGGTGATGTAGGTTTT 31
Db 747 CGTTATAGTTTTTCGGGACACGTGGGTTTT 718

Search completed: June 4, 2005, 12:20:07
Job time : 302.509 secs

This Page Blank (uspto)

went through one round of normalization, and was constructed by Bento Soares and M. Fatima Ronaldo.

ORIGIN

Constructed by Danto Boris and N. Rutana Donardo.

Query Match	72.9%;	Score	22.6;	DB	4;	Length	1317;
Best Local Similarity	86.2%;	Pred. No.	50;				
Matches	25;	Conservative	0;	Mismatches	4;	Indels	0; Gaps
							0;
Qy	2	CGTCTGGTCCTACGGGTGATGTAGGTTT	30				
Db	1031	CGTGCTGGTCCTTAAGGGTGATGTGGATT	1003				

```

ORIGIN
Query Match          71.0%; Score 22; DB 5; Length 692;
Best Local Similarity 83.3%; Pred. No. 84;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      2  CGTTCGTGCTTTACGGGTGATGTAGGTTTT 31
      |||||

```

```

ORIGIN
Query Match 68.4%; Score 21.2; DB 5; Length 474;
Best Local Similarity 88.5%; Pred. No. 1.8e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 TTCTGGTCTTACGGGTGATCTAGGTT 29
221 TGTGGTCTTACTGGGTGATCTGGTT 246

Db

```

Db 573 CTTTCTAGTCTTAAGGGAGCTGTAGGTTT 544

```

ORIGIN
Query Match 68.4%; Score 21.2; DB 5; Length 474;
Best Local Similarity 88.5%; Pred. NO. 1.8e+02;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 TTCTGCTCTTACGGGTGATGTAGGTT 29
221 TGTCTGCTCTTACGGGTGATGTAGGTT 246

Db

```

```

Seq primer: TR
Class: sheared ends.
Location/Qualifiers
  1..792
  /organism="Brassica oleracea"
  /mol_type="genomic DNA"
  /strain="TO1000DH3"
  /db_xref="taxon:3712"
  /clone_lib="BONLN49"
  /notes="Vector: pHOs1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOs1 using BstXI linkers"

ORIGIN
Query Match      66.5%; Score 20.6; DB 8; Length 792;
Best Local Similarity 85.2%; Pred. No. 3.4e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACGGGTGATCTAGG 27
    |||||
Db 511 ACCTTCTGGTCTTACGGGTGATCTAGG 537

RESULT 5
BZ449418
LOCUS
DEFINITION
  BZ449418 BO.1.6.2 KB tot Brassica oleracea genomic clone BZ449418,
  genomic survey sequence.
ACCESSION
  BZ449418
VERSION
  BZ449418.1 GI:26717819
KEYWORDS
  GSS.
SOURCE
  Brassica oleracea
  ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
    rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
  1 (bases 1 to 798)
  Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  Whole genome shotgun sequencing of Brassica oleracea
  Unpublished (2001)
  Other GSSs: BZ449418
  Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
  1..798
  /organism="Brassica oleracea"
  /mol_type="genomic DNA"
  /strain="TO1000DH3"
  /db_xref="taxon:3712"
  /clone_lib="BONRK40"
  /notes="Vector: pHOs1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOs1 using BstXI linkers"

ORIGIN
Query Match      66.5%; Score 20.6; DB 8; Length 798;
Best Local Similarity 85.2%; Pred. No. 3.5e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACGGGTGATCTAGG 27
    |||||
Db 287 ACCTTCTGGTCTTACGGGTGATCTAGG 313

RESULT 6
CR383334
LOCUS
DEFINITION
  CR383334 Bovine multi-stage muscles library (bcas) Bos taurus cDNA
  clone bcas0006a.e.08 3prim, mRNA sequence.
ACCESSION
  CR383334
VERSION
  CR383334.1 GI:47005238
KEYWORDS
  EST.
SOURCE
  Bos taurus (cow)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
    Bovinae; Bos.
    Chevalet,C.
    AGENAE, a French Animal Genome project
    Unpublished (2004)
    Contact: Martin P
    INRA, Genomique & Physiologie de la Lactation
    Domaine de Vilvert, 78352 JOUY-EN-JOSAS cedex, FRANCE
    Tel: +33 (0) 1.34.65.25.82
    Fax: +33 (0) 1.34.65.29.26
    Email: Patrice.Martin@jouy.inra.fr
    Sequence cleaned of vector, adaptor and repetitions. Contact us
    at signenasupport@jouy.inra.fr to obtain the chromatogram of this
    sequence.
    Plate: 0006 row: e column: 8.
    Location/Qualifiers
      1..344
      /organism="Bos taurus"
      /mol_type="mRNA"
      /db_xref="taxon:9913"
      /clone_lib="bcas0006a.e.08"
      /tissue_type="muscles : heart, longissimus thoracis,
      semitendinosus, masseter, cutaneous trunci"
      /dev_stages="from embryos to adults"
      /clone_lib="Bovine multi-stage muscles library (bcas)"
      /note="Clone distribution : AGENAE Resource centre.
      Francois PIUMI, Francois.Piumi@jouy.inra.fr, INRA, CEA
      Radiobiologie et Etude du genome (LREG), Domaine de
      Vilvert, 78352, Jouy-en-Josas cedex, FRANCE, +33 (0)
      1.34.65.28.02, +33 (0) 1.34.65.22.73"

ORIGIN
Query Match      65.8%; Score 20.4; DB 7; Length 344;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ACCTTCTGGTCTTACGGGTGATCTAGGTTT 30
    |||||
Db 245 ACTTCTGGTCAAGGGGTGATGAGGATT 274

RESULT 7
CR383335
LOCUS
DEFINITION
  CR383335 Bovine multi-stage muscles library (bcas) Bos taurus cDNA
  clone bcas0006a.e.08 5prim, mRNA sequence.
ACCESSION
  CR383335
VERSION
  CR383335.1 GI:47005239
KEYWORDS
  EST.
SOURCE
  Bos taurus (cow)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
    Bovinae; Bos.
    Chevalet,C.
    AGENAE, a French Animal Genome project
    Unpublished (2004)
    Contact: Martin P
    INRA, Genomique & Physiologie de la Lactation
    Domaine de Vilvert, 78352 JOUY-EN-JOSAS cedex, FRANCE
    Tel: +33 (0) 1.34.65.25.82
    Fax: +33 (0) 1.34.65.29.26

```


ACCESSION CC443427
VERSION CC443427.1 GI:30947492
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 928)
AUTHORS Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Reenick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bennetzen, J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PURGK01.7D
Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..928
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZM06.1.0.KB"
/notes="Vector: PCR4-TOPO; Site1: EcoRI; 0.6-1.0 kb high Cot selected genomic DNA library"
ORIGIN
Query Match 65.8%; Score 20.4; DB 8; Length 928;
Best Local Similarity 80.0%; Pred. No. 4.3e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 CGTCTGCTCTACGGGTGATGAGTTT 31
||||| ||||| ||||| ||||| |||||
DB 522 CCTCTGCTCTTTGGTGGTGATGATTTT 551
RESULT 18
LOCUS CA851238
DEFINITION D11F01_K13_11-ab1 cDNA Peking library 2, 4 day SCN3 Glycine max
ACCESSION CA851238
VERSION CA851238.1 GI:33388031
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 318)
AUTHORS Alkharouf, N.W., Khan, R. and Matthews, B.F.
TITLE Analysis of expressed sequence tags from roots of resistant soybean infected by the soybean cyst nematode
JOURNAL Unpublished (2002)
COMMENT Contact: Alkharouf, N.W.
Soybean Genomics and Improvement Laboratory (SGIL)
US Department of Agriculture (USDA), ARS, PSI
Bldg.006, Rm 118, 10300 Baltimore Ave., Beltsville, MD 20705-2350, USA
Tel: 301 504 5750
Fax: 301 504 5728
Email: alkharouf@ba.ars.usda.gov.
Location/Qualifiers
1..318
source

/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Peking"
/db_xref="taxon:3847"
/clone="D11F01"
/tissue_type="Roots"
/dev_stage="Seedlings"
/clone_lib="cDNA Peking library 2, 4 day SCN3"
/note="Vector: pBluescript SK-; cDNA clones from mRNA extracted from Peking roots 2 and 4 days past invasion."
ORIGIN
Query Match 64.5%; Score 20; DB 6; Length 318;
Best Local Similarity 82.1%; Pred. No. 5.6e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 TTCTGCTCTTACGGGTGATGAGTTT 31
||||| ||||| ||||| ||||| |||||
DB 105 TTCTGCTCTTTGGTGGTGATGAGTTT 132
RESULT 19
LOCUS BI943856
DEFINITION BI943856
sa45e07.Y1 Gm-cl004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl004-2269 5', mRNA sequence.
ACCESSION BI943856
VERSION BI943856.1 GI:16280235
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 380)
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpellding, J., Corvett, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this clone is listed in the 'Other ESTs on clone' field. This clone is available through: Biogenetic Services, 801 32nd Ave. Brookings, SD 57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..380
/organism="Glycine max"
/mol_type="mRNA"
/cultivar="Williams"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl004-2269"
/tissue_type="root"
/lab_host="XL10-Gold"
/clone_lib="Gm-cl004"
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; Root cDNA. The mRNA was isolated from entire roots of 8 day old 'Williams' seedlings which were propagated on paper towels with distilled water. Stragene's cDNA Synthesis kit (catalog #20401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated.
FEATURES
source


```

/db_xref="taxon:39947"
/clone="BR030042000 PLATE_A12_89_088.abl"
/tissue_type="roots"
/dev_stage="3-4 weeks"
/clone_lib="OA"
/notes="19 h 200mM NaCl"

ORIGIN
Query Match      64.5%; Score 20; DB 6; Length 621;
Best Local Similarity 82.1%; Pred. No. 6.2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 GTTCTGCTTACGGGTGATGTAGGTTT 30
||| ||||| ||||| ||||| |||||
Db 391 GTGCTGCTTATTGCTGGTCTAGGTTT 418

RESULT 25
CN225434/c
LOCUS WLA072A10.ab1 Wlbrain Gallus gallus cdna 5', mRNA linear EST 09-APR-2004
DEFINITION WLA072A10.ab1 Wlbrain Gallus gallus cdna 5', mRNA sequence.
ACCESSION CN225434
VERSION CN225434.1 GI:46328925
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 621)
Savolainen,P., Fitzsimmons,C.J., Arvestad,L., Andersson,L. and
Lundeberg,J.
EST analysis of brain and testis cdna libraries from White Leghorn
and Red Jungle Fowl
Unpublished (2004)
JOURNAL Department of Biotechnology
COMMENT Royal Institute of Technology, KTH
SE-106 91 Stockholm, SWEDEN
Tel: +46 (0)8 5537 8481
Fax: +46 (0)8 5537 8335
Email: Peter.Savolainen@biotech.kth.se
Seq primer: M13 reverse primer.
Location/Qualifiers
FEATURES
source
1..621
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn"
/db_xref="taxon:9031"
/sex="female"
/lab_host="ElectroMAX DH10B (Invitrogen)"
/clone_lib="Wlbrain"
/notes="Organ: brain; Vector: pSPORT-1; Site 1: Hind III;
Site 2: EcoRI; The cdna libraries were created with the
Superscript Plasmid System (Invitrogen)."
```

```

/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hisex"
/db_xref="taxon:9031"
/clone="ChST819ml1"
/dev_stage="22"
/lab_host="DH10B"
/clone_lib="CSEQCHN52"
/notes="Organ: limbs; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cdna
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cdna
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
```

```

ORIGIN
Query Match      64.5%; Score 20; DB 5; Length 706;
Best Local Similarity 82.1%; Pred. No. 6.3e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTACGGGTGATGTAGGTTT 31
||| ||||| ||||| ||||| |||||
Db 526 TTCTGCTCTCCGGGAGATTTTGGTTT 499

RESULT 27
BU120114/c
LOCUS 603141594F1 CSEQCHL16 Gallus gallus cdna clone ChST134k23 5', mRNA
DEFINITION 603141594F1 CSEQCHL16 Gallus gallus cdna clone ChST134k23 5', mRNA
sequence.
ACCESSION BU120114
VERSION BU120114.1 GI:25329297
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 815)
Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Boardman,P.E., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
Curt. Biol. 12 (22), 1965-1969 (2002)
1 (bases 1 to 706)
Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
A Comprehensive Collection of Chicken cDNAs
```

```

ORIGIN
Query Match      64.5%; Score 20; DB 7; Length 621;
Best Local Similarity 82.1%; Pred. No. 6.2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTTACGGGTGATGTAGGTTT 31
||| ||||| ||||| ||||| |||||
Db 527 TTCTGCTCTCCGGGAGATTTTGGTTT 500

RESULT 26
BU271768/c
LOCUS BU271768/c
DEFINITION 603819647F1 CSEQCHN52 Gallus gallus cdna clone ChST819ml1 5', mRNA
sequence.
ACCESSION BU271768
VERSION BU271768.1 GI:25542718
```

JOURNAL: Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1. .815
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton Line 151"
 /db_xref="taxon:9031"
 /clone="ChEST134k23"
 /sex="Female"
 /tissue_type="not cerebrum or cerebellum"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSECHL16"
 /notes="Organ: brain; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BsgI and BamHI sites [5'ggccgcgcgcagcccgccgcaaaaaag] [5'aattcttttttcggatccgggctgcacgc]"

ORIGIN
 Query Match 64.5%; Score 20; DB 5; Length 815;
 Best Local Similarity 82.1%; Pred. No. 6.4e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTCGTCTTACGGGTGATGTAGTGT 31
 Db 194 TTCTCGTCTTACGGGTGATGTAGTGT 167

RESULT 28
 BG616005/c
 LOCUS 602643446F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4774427 5',
 DEFINITION mRNA sequence.

ACCESSION BG616005
 VERSION BG616005.1 GI:13667376
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 827)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-f@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI647 row: 1 column: 12
 High quality sequence stop: 218.
 Location/Qualifiers
 1. .827

FEATURES
 source
 1. .827

JOURNAL: Curr. Biol. 12 (22), 1965-1969 (2002)
 MEDLINE 22335534
 PUBMED 12445392
 COMMENT Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES
 source
 1. .815
 Location/Qualifiers
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Compton Line 151"
 /db_xref="taxon:9031"
 /clone="ChEST134k23"
 /sex="Female"
 /tissue_type="not cerebrum or cerebellum"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSECHL16"
 /notes="Organ: brain; Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; Modification of pBluescript II KS(+) [Stratagene] vector to accommodate cDNA produced with the T-trimmed protocol (Construction of uni-directionally cloned cDNA libraries from messenger RNA for improved 3' end DNA sequencing by Glenn Fu, et al. U.S. Patent # 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI. Ligate in double stranded adaptor containing BsgI and BamHI sites [5'ggccgcgcgcagcccgccgcaaaaaag] [5'aattcttttttcggatccgggctgcacgc]"

ORIGIN
 Query Match 64.5%; Score 20; DB 5; Length 815;
 Best Local Similarity 82.1%; Pred. No. 6.4e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTCGTCTTACGGGTGATGTAGTGT 31
 Db 194 TTCTCGTCTTACGGGTGATGTAGTGT 167

RESULT 28
 BG616005/c
 LOCUS 602643446F1 NIH_MGC_61 Homo sapiens cDNA clone IMAGE:4774427 5',
 DEFINITION mRNA sequence.

ACCESSION BG616005
 VERSION BG616005.1 GI:13667376
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 827)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-f@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCMI647 row: 1 column: 12
 High quality sequence stop: 218.
 Location/Qualifiers
 1. .827

FEATURES
 source
 1. .827

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4774427"
 /tissue_type="embryonal carcinoma"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_61"
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccattggcc); Site_2: SfiI (ggccattggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGCGGACATG-DT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 clones contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 64.5%; Score 20; DB 4; Length 827;
 Best Local Similarity 82.1%; Pred. No. 6.4e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTCGTCTTACGGGTGATGTAGTGT 31
 Db 665 TTCTCGTCTTACGGGTGATGTAGTGT 638

RESULT 29
 CA994308
 LOCUS P02-16 Pst1-P fragment Gossypium hirsutum cDNA 5', mRNA
 DEFINITION EST.
 ACCESSION CA994308.1 GI:31074029
 VERSION EST.
 KEYWORDS Gossypium hirsutum (upland cotton)
 SOURCE Gossypium hirsutum
 ORGANISM Gossypium hirsutum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 1 (bases 1 to 864)
 Schulze, S.R., Bowers, J.E., Rong, J.K., Chapman, B.A., Abbey, C.,
 Brubaker, C.L., Chang, C., Garza, J.J., Park, C., Rainey, K.M.,
 Rastogi, V., Williams-Coplin, D., Zhao, X. and Paterson, A.H.
 A 3365 locus genetic recombination map of sequence-tagged sites
 reveals paleopolyploidy of diploid and tetraploid cotton
 (Gossypium)
 Unpublished (2002)
 Contact: Paterson, A.H.
 Plant Genome Mapping Laboratory
 University of Georgia, Center for Applied Genetic Technologies
 111 Riverbend Rd., Athens, GA 30602, USA
 Tel: 706 583 0162
 Fax: 706 583 0160
 Email: paterson@dogwood.botany.uga.edu
 Plate: CM11 row: D column: 05
 Seq primer: M13 Rev.

FEATURES
 Location/Qualifiers
 1. .864
 /organism="Gossypium hirsutum"
 /mol_type="mRNA"
 /cultivar="Accession 'TM-1'"
 /db_xref="taxon:3635"
 /tissue_type="Leaves"
 /lab_host="DH5-alpha"
 /clone_lib="Pst1-P fragment"
 /note="Vector: pUC(019); Site_1: Pst1"

ORIGIN
 Query Match 64.5%; Score 20; DB 6; Length 864;
 Best Local Similarity 82.1%; Pred. No. 6.4e+02;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 GTTCTGGTCTTACGGGTGATGAGTTT 30
 || ||||| ||||| ||||| ||||| |||||
 Db 755 GTACTGGTCTGACGGATGATTAAGTTT 782

RESULT 30
 BI852301/c
 LOCUS
 DEFINITION
 603378312f1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5391002 5',
 mRNA sequence.

ACCESSION BI852301
 VERSION BI852301.1 GI:15993048
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 936)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM1996 row: g column: 03
 High quality sequence stop: 727.

FEATURES

source
 1..936
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:5391002"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam2"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 64.5%; Score 20; DB 4; Length 936;
 Best Local Similarity 82.1%; Pred. No. 6.5e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 CGTCTGGTCTTACGGGTGATGAGTTT 29
 ||||| ||||| ||||| ||||| |||||
 Db 826 CGTCTGGTCTTACCGATGCTGTGGTT 799

RESULT 31
 CL114408/c
 LOCUS
 DEFINITION
 ISB1-60D2 Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-60D2,
 genomic survey sequence.

ACCESSION CL114408
 VERSION CL114408.1 GI:40608043
 KEYWORDS GSS.
 SOURCE xenopus tropicalis (western clawed frog)
 ORGANISM xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.

REFERENCE

AUTHORS
 1 (bases 1 to 1040)
 Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
 Mardis,E. and Wilson,R.
 TITLE
 JOURNAL
 COMMENT
 A physical map of the xenopus tropicalis genome
 Unpublished (2003)
 Contact: Richard K Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.edu
 Insert Length: 75000 Std Error: 0.00
 Seq primer: Sp6 ATTAGGTGACATATAG
 Class: BAC ends
 High quality sequence start: 7
 High quality sequence stop: 678.

FEATURES

source
 Location/Qualifiers
 1..1040
 /organism="Xenopus tropicalis"
 /mol_type="genomic DNA"
 /db_xref="taxon:8364"
 /clone="ISB1-60D2"
 /clone_lib="ISB1"
 /note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
 Library Segment 1"

ORIGIN

Query Match 64.5%; Score 20; DB 9; Length 1040;
 Best Local Similarity 82.1%; Pred. No. 6.6e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 GTTCTGGTCTTACGGGTGATGAGTTT 30
 ||||| ||||| ||||| ||||| |||||

Db 508 GTTATGGGCATACGTGTGATGAGCTT 481

RESULT 32

CC283372
 LOCUS
 DEFINITION
 CH261-128K21_RM1.1 CH261 Gallus gallus genomic clone CH261-128K21,
 genomic survey sequence.

ACCESSION CC283372
 VERSION CC283372.1 GI:30649812
 KEYWORDS GSS.
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus
 1 (bases 1 to 1067)
 Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
 Warren,W., Graves,T., Mardis,E. and Wilson,R.
 Gallus gallus BAC End Reads
 Unpublished (2003)
 Contact: Richard K. Wilson
 Genome Sequencing Center
 Washington University School of Medicine
 Email: submissions@wustl.edu
 Insert Length: 182000 Std Error: 0.00
 Seq primer: RM1 TACGACTCACTATAGGAGA
 Class: BAC ends
 High quality sequence start: 37
 High quality sequence stop: 825.

REFERENCE

AUTHORS
 1..1067
 Location/Qualifiers
 1..1067
 /organism="Gallus gallus"
 /mol_type="genomic DNA"
 /strain="Red Jungle Fowl"
 /db_xref="taxon:9031"
 /clone="CH261-128K21"
 /sex="female"
 /cell_line="UCD001, inbred 256"
 /clone_lib="CH261"
 /note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 CH261 Female Chicken library - for library and clone

FEATURES

source
 Location/Qualifiers
 1..1067

ordering information: <http://www.chori.org/bacpac>

```
ORIGIN
Query Match      64.5%; Score 20; DB 8; Length 1067;
Best Local Similarity 82.1%; Pred. No. 6.6e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGTTAGTTT 31
DB 337 TTCTCGTCTTCCGGGAGATTTTGGTTT 364

RESULT 33
CC219358      1255 bp      DNA      linear      GSS 12-MAY-2003
LOCUS CH261-69F20_Sp6.1 CH261 Gallus gallus genomic clone CH261-69F20,
DEFINITION genomic survey sequence.
ACCESSION CC219358
VERSION CC219358.1 GI:30541519
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE
AUTHORS Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Mardis, E. and Wilson, R.
TITLE Gallus gallus BAC End Reads
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 33
High quality sequence stop: 366.
Location/Qualifiers
1..1255
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-69F20"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"

ORIGIN
Query Match      64.5%; Score 20; DB 8; Length 1255;
Best Local Similarity 82.1%; Pred. No. 6.8e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TTCTGCTCTTACGGGTGATGTTAGTTT 31
DB 323 TTCTCGTCTTCCGGGAGATTTTGGTTT 350

RESULT 34
CD137210/c
LOCUS CD137210
DEFINITION MGI-0046U-A289-E02-U.G. mRNA sequence.
ACCESSION CD137210
VERSION CD137210.1 GI:34675102
KEYWORDS EST.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni

us-09-674-277-25.rst
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
REFERENCE
AUTHORS Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,
Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,
Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,
Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,
Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,
Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,
Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,
Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.
TITLE Transcriptome analysis of the acoelomate human parasite Schistosoma
mansoni
JOURNAL Nat. Genet. 35 (2), 148-157 (2003)
MEDLINE 228779926
PUBMED 12973350
COMMENT Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquímica
Instituto de Química - Universidade de São Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,
Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjowski@usp.br
This sequence was derived from the FAPESP Schistosoma mansoni EST
Genome Project. All sequences in the project were assembled and
annotated. This entry and all the assembled sequences can be seen
in the following URL http://bioinfo.iq.usp.br/schisto/
Plate: MGI-0046U-A289 row: 2 column: E.
Location/Qualifiers
1..215
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="MGI-0046U-A289-E02.G"
/sex="mixed pool"
/dev_stage="germball"
/lab_host="Biomphalaria glabrata"
/clone_lib="MGI-0046"
/note="Vector: pGEM T-easy"

ORIGIN
Query Match      63.9%; Score 19.8; DB 6; Length 215;
Best Local Similarity 77.4%; Pred. No. 6.5e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTTCTGCTTACGGGTGATGTTAGTTT 31
DB 173 ACAAATCCGGTCCGAGTGTAGTATTTT 143

RESULT 35
CD138405/c
LOCUS CD138405
DEFINITION MGI-0052T-D234-H09-U.B MGI-0052 Schistosoma mansoni cDNA clone
ACCESSION MGI-0052T-D234-H09.U.B, mRNA sequence.
VERSION CD138405
KEYWORDS EST.
SOURCE Schistosoma mansoni
ORGANISM Schistosoma mansoni

Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida; Schistosomatidae; Schistosomatidae; Schistosoma.
REFERENCE
AUTHORS Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,
Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,
Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,
Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,
Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,
Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,
Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,
```


directionally cloned at EcoRI and XhoI in Uni-ZAP XR. Following a single round of amplification phluescript sr phagemids were mass excised. A distribution channel for clones is being sought, but not currently available. Requests for clones cannot be honored."

ORIGIN

Query Match 63.9%; Score 19.8; DB 1; Length 416;
 Best Local Similarity 77.4%; Pred. No. 7.1e+02;
 Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ACGTCTGGCTCTACGGGTGATGATGTTT 31
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 226 ACGTCTGGCTCTTGGGGGTGTGTGAGTGT 196

RESULT 39
 CD134989

LOCUS
 DEFINITION MGI-0036U-A206-C12-U.B MGI-0036 Schistosoma mansoni cDNA clone
 MGI-0036U-A206-C12.B, mRNA sequence.

ACCESSION
 VERSION CD134989
 KEYWORDS EST.
 SOURCE Schistosoma mansoni
 ORGANISM Schistosoma mansoni
 Schistosoma mansoni
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
 Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.

REFERENCE
 1 (bases 1 to 431)
 Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M.,
 Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr.,
 Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.F.,
 Coulson, P.S., Dillon, G.P., Farias, L.P., Gregorio, S.P., Ho, P.L.,
 Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A.,
 Nascimento, A.L.T.O., Ohlweiler, F.P., Reis, E.M., Ribeiro, M.A.,
 Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kawano, T.,
 Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Menck, C.F.M.,
 Setubal, J.C., Leite, L.C.C. and Dias-Neto, E.
 Transcriptional analysis of the acelomate human parasite Schistosoma
 mansoni

TITLE
 Nat. Genet. 35 (2), 148-157 (2003)

JOURNAL
 MEDLINE 22879926
 PUBMED 12973350

COMMENT
 Contact: Dr. Sergio Verjovski-Almeida
 Departamento de Bioquímica
 Instituto de Química - Universidade de São Paulo
 Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP,
 Brasil
 Tel: +55-11-3091-2173
 Fax: +55-11-3091-2186
 Email: verj@iq.usp.br
 This sequence was derived from the FAPESP Schistosoma mansoni EST
 Genome Project. All sequences in the project were assembled and
 annotated. This entry and all the assembled sequences can be seen
 in the following URL http://bioinfo.iq.usp.br/schisto/
 Plate: MGI-0036U-A206 row: 12 column: C.

FEATURES
 Location/Qualifiers
 1..431
 /organism="Schistosoma mansoni"
 /mol_type="mRNA"
 /db_xref="taxon:6183"
 /clone="MGI-0036U-A206-C12.B"
 /sex="mixed pool"
 /dev_stage="germball"
 /lab_host="Biomphalaria glabrata"
 /clone_lib="MGI-0036"
 /note="vector: pGEM T-easy"

ORIGIN

Query Match 63.9%; Score 19.8; DB 6; Length 431;
 Best Local Similarity 77.4%; Pred. No. 7.2e+02;
 Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ACGTCTGGCTCTACGGGTGATGATGTTT 31
 ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 309 ACAATCGGTGCGAACCAGTGATGATGTTT 279

RESULT 38-
 AI946593/c

LOCUS
 DEFINITION AI946593 417 bp mRNA linear EST 08-JAN-2001
 ba28c03.y1 Drosophila melanogaster adult testis library Drosophila
 melanogaster cDNA clone ba28c03 5', mRNA sequence.

ACCESSION
 VERSION AI946593
 KEYWORDS EST.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
 1 (bases 1 to 417)
 Andrews, J., Bouffard, G.G., Cheadle, C., Lu, J., Becker, K.G. and
 Oliver, B.
 Gene discovery using computational and microarray analysis of
 transcription in the drosophila melanogaster testis

TITLE
 Genome Res. 10 (12), 2030-2043 (2000)

JOURNAL
 MEDLINE 20569492
 PUBMED 11116097

COMMENT
 On Aug 17, 1999 this sequence version replaced gi:5737019.
 Contact: Brian Oliver
 Laboratory of Cellular and Developmental Biology
 NIDDK, National Institutes of Health
 6 Center Drive MSC 2715, Bldg 6, Rm BI-13, Bethesda, MD 20892 USA
 Pax: (301) 496 5239
 Email: oliver@helix.nih.gov,
 http://www.niddk.nih.gov/intram/people/boliver.htm
 Tissue isolation and library construction performed at the National
 Institute of Diabetes and Digestive and Kidney Diseases, NIH (see
 http://www.niddk.nih.gov/intram/people/boliver.htm). DNA sequencing
 and analyses performed by National Institutes of Health Intramural
 Sequencing Center (NIISC; see http://www.nisc.nih.gov).
 Plate: 28 row: c column: 03
 Seq primer: M13RP1 reverse primer (ABI).
 Location/Qualifiers
 1..417
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="y[*] w[67c1]/y"
 /db_xref="taxon:7227"
 /clone="ba28c03"
 /sex="male"
 /dev_stage="1-5 day adult"
 /lab_host="Drosophila (Stratagene)"
 /clone_lib="Drosophila melanogaster adult testis library"
 /note="Organ: testis; Vector: pBluescript SK (Stratagene);
 Site 1: Eco I; Site 2: Xho I; Testes dissected from 1-5
 day adult y[*] w[67c1]/y males raised at 25°C. RNA
 isolated using Trizol (Life Technologies) and a single
 round of Poly(A)+ selection using Oligotex (Qiagen). cDNA
 library constructed using Stratagene ZAP-cDNA synthesis
 kit. Oligo dt-primed, size fractionated ~1-6 kb, and

Search completed: June 4, 2005, 11:46:46
Job time : 1895.65 secs

QY 1 ACGTTCGTGCTTACGGGTGATGATGTTT 31
||| ||||| ||| ||||| ||||| |||||
DB 251 ACAATCCGTCGAACGAGTGATGATGTTT 281

RESULT 40

CV253036/c

LOCUS

CV253036/c

DEFINITION

cDNA clone PX0019_P03 3', mRNA sequence.

ACCESSION

CV253036

VERSION

CV253036.1

KEYWORDS

EST.

SOURCE

Populus balsamifera subsp. trichocarpa (Populus trichocarpa)

ORGANISM

Populus balsamifera subsp. trichocarpa

REFERENCE

1 (bases 1 to 481)

AUTHORS

Ralph, S., Cooper, D., Kolosova, N., Oddy, C., Butterfield, Y., Kirpatrick, R., Liu, J., Palmquist, D., Stott, J., Barber, S., Yang, G., Babakaiff, R., Brown-John, M., Chand, S., Featherstone, R., Masson, A., Mayo, M., Moran, J., Olson, T., Wong, D., Ritland, C.E., Siddiqui, A., Holt, R., Jones, S., Marra, M., Ellis, B.E., Douglas, C., Ritland, K. and Bohlmann, J.

TITLE

The poplar transcriptome: Analysis of expressed sequence tags from multiple cDNA libraries

JOURNAL

Unpublished (2004)

COMMENT

Contact: Joerg Bohlmann

Genome BC forest genomics program

University of British Columbia

UBC Biotechnology Laboratory, 6174 University Boulevard, Rm. 237,

Vancouver, British Columbia, Canada, V6T 1Z3

Tel: 1-604-822-0282

Fax: 1-604-822-6097

Email: bohlmann@interchange.ubc.ca

Plate: PX0019 row: P column: 03

High quality sequence stop: 506

POLYA=Yes.

FEATURES

Location/Qualifiers

1..481

/organism="Populus balsamifera subsp. trichocarpa"

/mol_type="mRNA"

/cultivar="383-2499 (Nisqually-1)"

/sub_species="trichocarpa"

/db_xref="taxon:3694"

/clone="PX0019_P03"

/sex="Female"

/lab_host="E. coli DH10B T1 phage resistant cells"

/clone_lib="PT-X-FL-A-1"

/note="Vector: pBluescript II SK (+) XR; Site 1: SstI (5'

end of cDNA); Site 2: XhoI (3' end of cDNA); Outer xylem

from 8 year old trees harvested within the Boise Cascade

region of Washington state on May 15th, 2001. cDNA was

prepared from 20 micrograms of mRNA according to the

full-length cDNA library construction method described by

Carninci P. et al. (2000), Genome Research

10(10):1617-1630 and directionally ligated into the

pBluescript II SK (+) XR vector digested with SstI (5'

end) and XhoI (3'). Plasmid DNA was then transformed by

electroporation into DH10B cells (Invitrogen) for

propagation."

ORIGIN

Query Match

Best Local Similarity 63.9%; Score 19.8; DB 7; Length 481;

Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 ACGTTCGTGCTTACGGGTGATGATGTTT 31

||| ||||| ||| ||||| ||||| |||||

DB 108 ATGGTCTGCTTTTACAAGGATGATGTTT 78

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 04:22:34 ; Search time 234.976 Seconds
(without alignments)
780.981 Million cell updates/sec

Title: US-09-674-277-26

Perfect score: 31

Sequence: 1 tagtgaacggtgacagcatatcagacggct 31

Scoring table: IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	31	Aaz36126	Primer de
2	31	100.0	1181	3 AAZ36102	Nucleic a
3	20	64.5	918	10 ADD16751	ADD16751 DNA (Seqi
4	19.8	63.9	1836	4 ABL17955	ABL17955 Drosophil
5	19.8	63.9	3917	4 ABL17954	ABL17954 Drosophil
6	19.4	62.6	1470	13 ADT48557	ADT48557 Bacterial
7	19.4	62.6	2367	4 ABL22265	ABL22265 Drosophil
8	19.4	62.6	2367	12 AD007755	Ado07755 Fly polyn
9	19.4	62.6	2459	8 ABV75073	Abv75073 Drosophil
10	19.4	62.6	10386	4 ABL22264	ABL22264 Drosophil
11	19	61.3	3694	4 AAS26150	Aas26150 Human cdn
12	19	61.3	3694	8 ABX73491	Abx73491 Human nov
13	19	61.3	71678	11 ACN44284	Acn44284 Mouse gen
14	18.8	60.6	236	3 AAC17472	Aac17472 Human sec
15	18.8	60.6	948	10 ADE07620	Ado07620 Novel cod
16	18.8	60.6	1437	11 ABD10016	Abd10016 Pseudomon
17	18.8	60.6	1465	6 ABL58968	ABL58968 Human tum
18	18.8	60.6	1506	11 ABD09998	Abd09998 Pseudomon
19	18.8	60.6	1554	8 ACCT79230	Acc79230 Human INP
20	18.8	60.6	1626	11 ABD10038	Abd10038 Pseudomon

21	18.8	60.6	1653	5 AAh65032	Aah65032 C glutami
22	18.8	60.6	1653	8 ACA01840	Aca01840 C. glutam
23	18.8	60.6	2324	4 AAK94137	Aak94137 Human ful
24	18.8	60.6	2324	12 ADL30609	Adl30609 Full leng
25	18.8	60.6	2640	6 ABL61989	AbL61989 Colon ade
26	18.8	60.6	2640	12 ADN03923	Adn03923 Antipsori
27	18.8	60.6	2640	12 ADP10518	Adp10518 Reference
28	18.8	60.6	4609	4 AAK83787	Aak83787 Human imm
29	18.8	60.6	270150	11 ADP65796	Adp65796 Human top
30	18.8	60.6	349980	5 AAh64966	Aah64966 C glutami
31	18.6	60.0	1398	5 AAS68064	Aas68064 DNA encod
32	18.6	60.0	1464	6 ABN68834	Abn68834 Streptoco
33	18.6	60.0	1467	8 ACA50653	Acas0663 Prokaryot
34	18.4	59.4	251	3 AAC09705	Aac09705 Human sec
35	18.4	59.4	438	5 ABV38251	Abv38251 Human pro
36	18.4	59.4	485	5 ABV24200	Abv24200 Human pro
37	18.4	59.4	1086	13 ADS56949	Ads56949 Bacteri
38	18.4	59.4	1289	5 ABV25356	Abv25356 Human pro
39	18.4	59.4	1289	5 ABV25037	Abv25037 Human T c
40	18.4	59.4	1674	6 AAD24919	Aad24919 Human T c
41	18.4	59.4	1799	6 ABZ35381	Abz35381 Human gen
42	18.4	59.4	2117	3 AAC37195	Aac37195 Arabidops
43	18.4	59.4	2117	3 AAC47417	Aac47417 Arabidops
44	18.4	59.4	2163	10 ADE07608	Ade07608 Novel cod
45	18.4	59.4	2243	12 ADO63638	Ado63638 Transcrip

ALIGNMENTS

RESULT 1
AAZ36126
ID AAZ36126 standard; DNA; 31 BP.

AC AAZ36126;

DT 11-FEB-2000 (first entry)

XX Primer derived from a nucleic acid sequence specific to EHEC.

DE Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;

KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;

KW PCR primer; probe; ss.

XX Synthetic.

OS Escherichia coli.

XX WO9955908-A2.

PN 04-NOV-1999.

PD 27-APR-1999; 99WO-FR001000.

XX 28-APR-1998; 98FR-00005329.

XX (SNFI) PASTEUR SANOFI DIAGNOSTICS.

PA Frechon DTM, Laure FC, Thierry D;

PI WPI; 2000-013443/01.

DR New nucleic acid containing sequences specific to enterohemorrhagic

XX Escherichia coli, particularly serotype O157:H7, used for detecting these

PT bacteria in food.

XX Claim 5; Page 27; 48pp; French.

CC AAZ36103-27 represent fragments derived from nucleic acid sequences

CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are

CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to

CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and

CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).

CC The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies

XX SQ Sequence 31 BP; 9 A; 6 C; 10 G; 6 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 31;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
|||||
DB 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31

RESULT 2

AAZ36102
ID AAZ36102 standard; DNA; 1181 BP.

XX AC AAZ36102;

XX DT 11-FEB-2000 (first entry)

XX DE Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX KW Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

XX KW enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.

XX OS Escherichia coli.

XX PN WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX PS WPI; 2000-013443/01.

XX PT New nucleic acid containing sequences specific to enterohemorrhagic
XX PT Escherichia coli, particularly serotype O157:H7, used for detecting these
XX PT bacteria in food.

XX PS Claim 1; Fig 2; 48pp; French.

XX CC The present sequence is specific to enterohemorrhagic Escherichia coli
XX CC (EHEC). The sequence associated with the presence of virulence factors
XX CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
XX CC 68% homology with the virK gene which codes for virulence proteins of
XX CC Shigella flexneri. The present sequence is of plasmid origin. Fragments
XX CC of the present sequence are used, as probes and primers, for detection of
XX CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
XX CC animal samples, foods or the environment. The fragments are also useful
XX CC for epidemiological studies

XX SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match 100.0%; Score 31; DB 3; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31

DB 970 TAGTGAAGCGGTGACAGCATATCAGACGGCT 1000

RESULT 3

ADD16751

ID ADD16751 standard; DNA; 918 BP.

XX AC ADD16751;

XX DT 15-JAN-2004 (first entry)

XX DE DNA (SeqID 819) that confers an altered visual phenotype in plants.

XX KW ds; visual phenotype; plant; architecture; leaf surface; chlorotic;

XX KW bleaching; etching; wet leaf; stunting; elongation; texture;

XX KW agronomic trait; growth regulation; dwarf variety; insect resistance;

XX KW heat stress; transgenic.

XX OS Unidentified.

XX PN WO2003020741-A1.

XX PD 13-MAR-2003.

XX PF 30-AUG-2002; 2002WO-US027880.

XX PR 31-AUG-2001; 2001US-0316326P.

XX PA (DOWC) DOW CHEM CO.

XX PA (DOWC) DOW AGROSCIENCES LLC.

XX PI Crosley R, Skokut T, Ruegger M, Larrinua I, Shukla V;

XX PS WPI; 2003-300858/29.

XX PT Novel isolated nucleic acid derived from Nicotiana benthamiana, Oryzae
XX PT sativa, Saccharomyces cerevisiae, and Papaver rhoeas, useful for
XX PT conferring altered visual phenotypes in plants.

XX PS Claim 1; SEQ ID NO 819; 517pp; English.

XX CC This invention relates to the identification and isolation of novel
XX CC nucleic acid molecules that confer altered visual phenotypes in plants.
XX CC Specifically, it refers to modifications of plant architecture and/or
XX CC leaf surface features in plants, such as chlorotic, bleaching, etching,
XX CC wet leaf, stunting, elongation and texture phenotypes, which are thought
XX CC will be agronomic traits beneficial to the farmer. As such, these novel
XX CC phenotypes can affect growth regulation i.e. useful for creating dwarf
XX CC varieties, exhibit resistance to insects or heat stress, confer changes
XX CC in pigment content such that plants have enhanced vitamin production or
XX CC delayed senescence and also for example produce plants that control the
XX CC production of ethylene. Furthermore, the present invention comprises
XX CC generating transgenic plants, as well as reproducibly altering the visual
XX CC phenotype of plant seeds, plant tissues and plant cells containing the
XX CC polynucleotides described herein. This polynucleotide is a homologue of a
XX CC DNA sequence that confers an altered visual phenotype when expressed in
XX CC plants, the method of the invention.

XX SQ Sequence 918 BP; 211 A; 241 C; 243 G; 223 T; 0 U; 0 Other;

Query Match 64.5%; Score 20; DB 10; Length 918;
Best Local Similarity 82.1%; Pred. No. 59;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TGAAGCGGTGACAGCATATCAGACGGCT 31

DB 636 TGTAGCGCTGACAGGATCTCAGAGGCT 663

RESULT 4

ABL17955

ID ABL17955 standard; DNA; 1836 BP.

XX AC ABL17955;

XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 5338.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 5338; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1836 BP; 427 A; 486 C; 526 G; 397 T; 0 U; 0 Other;
XX
XX Query Match 63.9%; Score 19.8; DB 4; Length 1836;
XX Best Local Similarity 91.3%; Pred. No. 83;
XX Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 3 GTGAGCGGTGACGCAATATCAG 25
XX Db 968 GTGAGCGGTGACGCAATCAG 990
XX
XX RESULT 5
XX ID ABL17954/c
XX AC ABL17954;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 5335.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX DR WPI; 2001-656860/75.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions.
XX PS Claim 1; SEQ ID NO 5335; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
XX CC ABB72072). The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 3917 BP; 1000 A; 958 C; 942 G; 1017 T; 0 U; 0 Other;
XX
XX Query Match 63.9%; Score 19.8; DB 4; Length 3917;
XX Best Local Similarity 91.3%; Pred. No. 95;
XX Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 3 GTGAGCGGTGACGCAATATCAG 25
XX Db 1869 GTGAGCGGTGACGCAATCAG 1847
XX
XX RESULT 6
XX ID ADT48557/c
XX AC ADT48557;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polynucleotide #23308.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polynucleotide; gene; ss.
XX OS Bacteria.
XX PN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 46995; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1470 BP; 394 A; 342 C; 377 G; 357 T; 0 U; 0 Other;

Query Match 62.6%; Score 19.4; DB 13; Length 1470;
Best Local Similarity 95.2%; Pred. No. 1.2e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 GTGACAGCATATCAGACGGCT 31
|||||
Db 208 GTGACAGCAATCAGACGGCT 188

RESULT 7
ABL22265
ID ABL22265 standard; DNA; 2367 BP.
XX
AC ABL22265;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 18268.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX

DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 18268; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2367 BP; 506 A; 690 C; 648 G; 523 T; 0 U; 0 Other;

Query Match 62.6%; Score 19.4; DB 4; Length 2367;
Best Local Similarity 79.3%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGCACGACATATCAGACGGC 30
|||||
Db 1315 AGTGAAGCGCTGGCAGCATTTGCGGCGGC 1343

RESULT 8
ADO07755
ID ADO07755 standard; cDNA; 2367 BP.
XX
AC ADO07755;
XX
DT 01-JUL-2004 (first entry)
XX
DE Fly polynucleotide #36.
XX
KW Fly; gene; ss; fat cell number; fat cell size; obesity; diabetes;
KW anorectic; antidiabetic.
XX
OS Diptera.
XX
FN US2004071700-A1.
XX
PD 15-APR-2004.
XX
PF 09-OCT-2002; 2002US-00267502.
XX
PR 09-OCT-2002; 2002US-00267502.
XX
PA (LIFE-) LIFE SCI DEV CORP.
XX
PI Kim J, Galant R;
XX
DR WPI; 2004-328526/30.
DR P-PSDB; ADO07972.
XX
PT Identifying compounds that influence fat cell number or size for treating
PT or preventing obesity or diabetes by exposing the cell to the agent and
PT identifying fat cell number or size relative to cells not exposed to the
PT agent.
XX
PS Claim 1; SEQ ID NO 81; 275pp; English.
XX
CC The invention relates to a method of identifying compounds that influence
CC fat cell number or size comprising providing a cell that expresses a gene
CC and an agent, exposing the cell to the agent and identifying fat cell
CC number or size relative to cells not exposed to the agent. The method
CC also comprises providing an expression vector and an agent, exposing the
CC vector to the agent, detecting a change in expression of the gene

CC relative to expression of the gene in an expression vector not exposed to
CC the agent, treating a subject with the agent and identifying fat cell
CC number or size in the subject. The agent comprises an antisense
CC oligonucleotide. The subject comprises a mammal, preferably a human. The
CC method also comprises providing a polypeptide and an agent, exposing the
CC polypeptide to the agent, detecting binding of the agent to the
CC polypeptide or a change in an activity of the polypeptide, treating a
CC subject with the agent and identifying fat cell number or size in the
CC subject. The agent comprises an antibody. A method of regulating fat cell
CC number or size comprises providing a subject containing fat cells and an
CC agent that changes the expression of a gene, and treating the subject
CC with the agent under conditions so that fat cell size or number in the
CC subject is altered. The method is useful for identifying compounds that
CC influence fat cell number or size, for preparing a composition for
CC treating or preventing obesity or diabetes. This sequence represents fly
CC cDNA used in the scope of the invention.

XX Sequence 2367 BP; 506 A; 690 C; 648 G; 523 T; 0 U; 0 Other;

Query Match 62.6%; Score 19.4; DB 12; Length 2367;
Best Local Similarity 79.3%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACACATATCAGACGGC 30
Db 1315 AGTGAAGCGGTGACACATATTCGCGCGC 1343

RESULT 9
ABV75073

ID ABV75073 standard; cDNA; 2459 BP.

AC ABV75073;

DT 19-FEB-2003 (first entry)

DE Drosophila DevG22 protein encoding cDNA.

XX Protein disulfide isomerase; DevG20; ABC transporter; DevG4; DevG22;
KW anorectic; immunomodulator; antidepressant; antidiabetic; hypotensive;
KW antiarteriosclerotic; antilipemic; osteopathic; antiarthritic; gene;
KW litholytic; hepatotropic; cytostatic; neuroprotective; gene therapy;
KW transgenic; ss.

OS Drosophila sp.

FH Key Location/Qualifiers

FT CDS 576..2459

FT /*tag= a

FT /product= "DevG22 protein"

XX WO200279238-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-EP003540.

XX 02-APR-2001; 2001EP-00108315.

XX 01-JUN-2001; 2001EP-00113419.

XX (DEV-) DEVELOPEN ENTWICKLUNGSBIOLOGISCHE FORSCH.

XX Eulenber K, Broenner G, Ciossek T, Haeder T, Steuernagel A;

XX WPI; 2003-067420/06.

XX P-PSDB; ABB82646.

XX New protein disulfide isomerase and ABC transporter nucleic acids or
PT polypeptides, useful for diagnosing, treating or preventing obesity,
PT hypertension, heart disease, hypercholesterolemia, osteoarthritis,
PT gallstones or cancer.

XX Example 8; Fig 15A; 99pp; English.

XX The invention relates to pharmaceutical composition comprising carriers,
CC diluents and/or adjuvants, with any of: (a) a nucleic acid molecule of
CC the protein disulfide isomerase (DevG20) or ABC transporter (DevG4) or
CC DevG22) gene family; (b) a polypeptide encoded by (a); (c) a fragment or
CC variant of (a) or (b); or (d) an antibody, an aptamer or another receptor
CC recognizing (a) or (b). The composition is useful for manufacturing an
CC agent for detecting and/or verifying, diagnosing, treating, alleviating
CC or preventing a metabolic disorder, e.g. obesity, adipositas, eating/body
CC weight disorders, cachexia (wasting), pancreatic dysfunction (diabetes
CC mellitus), hypertension, arteriosclerosis, coronary artery disease (CAD),
CC coronary heart disease, hypercholesterolemia, dyslipidemia,
CC osteoarthritis, gallstones, cancer (cancer of the reproductive organs),
CC sleep apnea, disorders related to ROS production and neurodegenerative,
CC diseases in cells, cell masses, organs and/or subjects. The inhibitors,
CC modulators or agents identified above are useful for treating, preventing
CC or alleviating the diseases mentioned. The nucleic acid molecule of
CC DevG20, DevG4 and/or Dev is also useful for preparing a non-human animal
CC which over- or underexpresses the DevG20, DevG4 and/or Dev gene product.
CC The present sequence represents a Drosophila DevG22 protein encoding cDNA

XX Sequence 2459 BP; 624 A; 656 C; 649 G; 530 T; 0 U; 0 Other;

Query Match 62.6%; Score 19.4; DB 8; Length 2459;
Best Local Similarity 79.3%; Pred. No. 1.3e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACACATATCAGACGGC 30

Db 1890 AGTGAAGCGGTGACACATATTCGCGCGC 1918

RESULT 10
ABL22264/c

ID ABL22264 standard; DNA; 10386 BP.

XX ABL22264;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 18265.

XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 18265; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 10386 BP; 3035 A; 2174 C; 2318 G; 2859 T; 0 U; 0 Other;
Query Match 62.6%; Score 19.4; DB 4; Length 10386;
Best Local Similarity 79.3%; Pred. No. 1.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACGACATATCAGACGGC 30
Db 5336 AGTGAAGCGGTGACGACATTTGCGCGGC 5308
RESULT 11
AAS26150
ID AAS26150 standard; cDNA; 3694 BP.
XX
AC AAS26150;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human cDNA encoding a novel secreted protein, Seq ID 329.
XX
KW Human; immunosuppressive; antiarthritic; ss; antirheumatic; cytostatic;
KW cardiant; vasotropic; cerebroprotective; nontropic; neuroprotective;
KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;
KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;
KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;
KW cerebral ischaemia; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; epithelial cell proliferation; skin ageing; food additive;
KW preservative; antiproliferative.
XX
OS Homo sapiens.
XX
PN WO200155322-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001341.
XX
31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.


```

PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
(ROSE/) ROSEN C A.
(RUBE/) RUBEN S M.
(BARA/) BARASH S C.
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-147444/14.
DR P-PSDB; ABU55231.
XX
XX New polypeptides and nucleic acids, useful in gene therapy for treating,
PT inhibiting or preventing e.g. neural, immune system, muscular,
PT respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT renal disorders.
XX
XX Claim 1; SEQ ID NO 329; 402pp; English.
XX
XX The invention relates to human novel polypeptides and their associated
CC polynucleotides. The polypeptides and polynucleotides are useful in gene
CC therapy for treating, inhibiting or preventing neural disorders, immune
CC system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC (e.g. congenital heart defects, Ebstein's anomaly and hypoplastic left
CC heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC infarction) and cancerous diseases. Sequences ABX73173-ABX74167 represent
CC human novel polynucleotides of the invention
XX
XX Sequence 3694 BP; 1103 A; 934 C; 839 G; 817 T; 0 U; 1 Other;
XX
Query Match 61.3%; Score 19; DB 8; Length 3694;
Best Local Similarity 75.9%; Pred. No. 2.2e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
XX
QY 2 AGTGAAGCGGTGACAGCATATCAGAGCGC 30
Db 1613 AGTGAAGCTGGAAACGCATATCAGAGTGC 1641
XX
RESULT 13
ACN44284
ID ACN44284 standard; DNA; 71678 BP.
XX
XX ACN44284;
AC
XX
XX 18-NOV-2004 (first entry)
DT
XX
XX Mouse genomic sequence mCG16994.
XX
XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX
XX Mus musculus.
OS
XX
XX WO2003073826-A2.
PN
XX

```

```

PD 12-SEP-2003.
XX
XX 28-FEB-2003; 2003WO-US006235.
XX
XX 01-MAR-2002; 2002US-00087192.
XX
XX (SAGR-) SAGRES DISCOVERY.
PA
XX Morris DW;
PI
XX WPI; 2003-328604/31.
DR
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX
XX Claim 1; SEQ ID NO 655; Opp; English.
XX
XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX
XX Sequence 71678 BP; 20309 A; 14954 C; 15127 G; 20103 T; 0 U; 1185 Other;
XX
Query Match 61.3%; Score 19; DB 11; Length 71678;
Best Local Similarity 81.5%; Pred. No. 3.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
XX
QY 2 AGTGAAGCGGTGACAGCATATCAGAGC 28
Db 33654 AGTGTAGCAGTGACAGCATCTCAAAAG 33680
XX
RESULT 14
AAC17472
ID AAC17472 standard; cDNA; 236 BP.
XX
XX AAC17472;
AC
XX
XX 06-OCT-2000 (first entry)
DT
XX
XX Human secreted protein 5' EST, SEQ ID NO: 21547.
DE
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1033401-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 21-FEB-2000; 2000EP-00200610.
PF
XX
XX 26-FEB-1999; 99US-0122487P.
PR
XX
XX (GEST ) GENSET.
PA
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX

```



```
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 1437 BP; 206 A; 498 G; 482 C; 251 T; 0 U; 0 Other;

Query Match          60.6%; Score 18.8; DB 11; Length 1437;
Best Local Similarity 76.7%; Pred. No. 2.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGC 30
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 318 TCGTGAAGCTGGGCCAGCATCTCAGCCTGC 347

RESULT 17
ABL58968
ID ABL58968 standard; cDNA; 1465 BP.
XX
AC ABL58968;
XX
DT 19-JUL-2002 (first entry)
XX
DE Human tumour marker cDNA Li9-1.
XX
KW Human; tumour; cytostatic; cutaneous T cell lymphoma; CTCL; vaccine;
KW antigen-presenting cell; tumour-specific T cell; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200238803-A2.
XX
PD 16-MAY-2002.
XX
PF 08-NOV-2001; 2001WO-DE004229.
XX
PR 08-NOV-2000; 2000DE-01055285.
XX
PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Bichmuller S, Schadendorf D, Usener D;
DR WPI; 2002-426959/45.
DR P-PSDB; ABB77441.
XX
PT Composition containing tumor-associated nucleic acid, useful for
PT diagnosis and treatment of tumors, especially cutaneous T cell lymphoma,
PT also derived proteins and antibodies.
XX
PS Claim 1; Fig 18; 84pp; German.
XX
CC The invention relates to a diagnostic composition containing at least one
CC of 23 nucleotide sequences (I, ABL58901-ABL58950) with altered expression
CC associated with tumors. (I), including antisense sequences and
CC ribozymes, also proteins (II, ABB77424-ABB77445) encoded by them and
CC antibodies specific for (II), are useful for diagnosis, monitoring and
CC treatment of tumors, especially cutaneous T cell lymphoma (CTCL). (II)
CC and antibodies to (II) are useful for vaccination. (II) can also be used
CC to prepare pre-loaded antigen-presenting cells or tumour-specific T cells
XX
SQ Sequence 1465 BP; 334 A; 423 C; 443 G; 265 T; 0 U; 0 Other;

Query Match          60.6%; Score 18.8; DB 6; Length 1465;
Best Local Similarity 76.7%; Pred. No. 2.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGGCT 31
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 493 ACTGATGCGATGAAGAATATAACAGGGCT 522

RESULT 18
ABD09998/c
ID ABD09998 standard; DNA; 1506 BP.
XX
AC ABD09998;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polynucleotide #8602.
XX
KW Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
DR WPI; 2003-615309/58.
DR P-PSDB; ABO76427.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 8602; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABD01397-
CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 1506 BP; 256 A; 511 C; 518 G; 221 T; 0 U; 0 Other;

Query Match          60.6%; Score 18.8; DB 11; Length 1506;
Best Local Similarity 76.7%; Pred. No. 2.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGC 30
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 905 TCGTGAAGCTGGGCCAGCATCTCAGCCTGC 876

RESULT 19
ACC79230
ID ACC79230 standard; cDNA; 1554 BP.
XX
AC ACC79230;
XX
DT 31-JUL-2003 (first entry)
XX
DE Human INPP5A encoding cDNA SEQ ID NO:1.
XX
```

KW Human; MSREBP; sterol regulatory element binding protein; SREBP;
KW antiarteriosclerotic; antilipemic; antidiabetic; cardiant; diabetes;
KW atherosclerosis; hyperlipidaemia; cardiovascular disease; gene; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 95..1333
FT /*tag= a
FT /product= "INPP5A"
XX WO2003033656-A2.
XX PN
XX PD
XX PF
XX PF 15-OCT-2002; 2002WO-US032807.
XX PR 16-OCT-2001; 2001US-0329890P.
XX PR 21-DEC-2001; 2001US-0342287P.
XX PR 21-DEC-2001; 2001US-0342288P.
XX PR 14-JAN-2002; 2002US-0348096P.
XX PR 25-JAN-2002; 2002US-0351361P.
XX PR 25-JAN-2002; 2002US-0351401P.
XX PR 25-JAN-2002; 2002US-0351403P.
XX PR 21-FEB-2002; 2002US-0358826P.
XX PR 27-MAR-2002; 2002US-0368615P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Kadyk LC, O'brien CL, Kong EC, Hammonds GR;
XX WPI; 2003-381794/36.
XX DR P-PSDB; ABR65950.
XX PT Identifying a candidate sterol regulatory element binding protein (SREBP)
XX PT pathway modulating agent for treating e.g. hyperlipidemia, by contacting
XX PT an assay system comprising an MSREBP polypeptide or nucleic acid with a
XX PT test agent.
XX PS Example; Page 50-51; 120pp; English.
XX CC The present invention describes a method for identifying a candidate
XX CC sterol regulatory element binding protein (SREBP) pathway modulating
XX CC agent. The method comprises: (a) providing an assay system comprising an
XX CC MSREBP (human SREBP) polypeptide or nucleic acid; (b) contacting the
XX CC assay system with a test agent under conditions where the system provides
XX CC a reference activity in the presence of the test agent; and (c) detecting
XX CC a test agent-biased activity of the screening assay system, where a
XX CC difference between the test-biased activity and the reference activity
XX CC identifies the test agent as a candidate SREBP pathway modulating agent.
XX CC Also described is a method for modulating SREBP pathway activity in a
XX CC mammalian cell by contacting the cell with an agent that specifically,
XX CC binds an MSREBP polypeptide or nucleic acid. SREBP has
XX CC antiarteriosclerotic, antilipemic, antidiabetic and cardiant activities.
XX CC The method is useful for identifying modulators of SREBP pathway for
XX CC treating atherosclerosis, hyperlipidaemia, diabetes, or cardiovascular
XX CC diseases. The present sequence encodes a human MSREBP designated INPP5A,
XX CC which is used in an example from the present invention
XX SQ Sequence 1554 BP; 367 A; 444 C; 422 G; 321 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 8; Length 1554;
Best Local Similarity 76.7%; Pred. No. 2.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACAGCATATCAGCGCT 31
Db 326 AGTGATCCGATGAAGAATATAACAGCGCT 355
RESULT 20
ABD10038
ID ABD10038 standard; DNA; 1626 BP.

XX ABD10038;
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polynucleotide #8642.
XX Bacterial infection; gene; ds; Pseudomonas aeruginosa infection;
KW antibacterial.
XX OS Pseudomonas aeruginosa.
XX PN US6551795-B1.
XX PD 22-APR-2003.
XX PF 18-FEB-1999; 99US-00252991.
XX PR 18-FEB-1998; 98US-0074788P.
XX PR 27-JUL-1998; 98US-0094190P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX WPI; 2003-615309/58.
XX DR P-PSDB; ABO76467.
XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
XX PT pathological conditions resulting from bacterial infection.
XX PS Disclosure; SEQ ID NO 8642; 455pp; English.
XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
XX CC polynucleotides encoding them. The sequences are useful in diagnosis and
XX CC therapy of pathological conditions, as molecular targets for diagnostics,
XX CC prophylaxis and treatment of pathological conditions resulting from a
XX CC bacterial infection, for evaluating a compound, such as a polypeptide,
XX CC for the ability to bind a P. aeruginosa nucleic acid, as components of
XX CC effective antibacterial targets, as targets for antibacterial drugs,
XX CC including anti-P. aeruginosa drugs, as templates for recombinant
XX CC production of P. aeruginosa-derived peptides or polypeptides, as target
XX CC components for diagnosis and/or treatment of P. aeruginosa-caused
XX CC infection, and in detection of P. aeruginosa sequences or other sequences
XX CC of Pseudomonas species using biochip technology. Sequences ABD01397-
XX CC ABD17967 represent P. aeruginosa polynucleotides of the invention. Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format from USPTO at
XX CC seqdata.uspto.gov/sequence.html
XX SQ Sequence 1626 BP; 221 A; 564 C; 553 G; 288 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 11; Length 1626;
Best Local Similarity 76.7%; Pred. No. 2.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGC 30
Db 209 TCGTGAAGCTGGGCCAGCATCTCAGCGCTGC 238
RESULT 21
AAH65032
ID AAH65032 standard; DNA; 1653 BP.
XX AC AAH65032;
XX 26-SEP-2001 (first entry)
XX C glutamicum coding sequence fragment SEQ ID NO: 67.
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis; ds.
XX Corynebacterium glutamicum.
XX EP1108790-A2.
XX 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
XX 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Nakagawa S, Mizoguchi H, Ando S, Hayaishi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
DR P-PSDB; AAG89813.
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX Claim 8; SEQ ID NO 67; 246pp + Sequence Listing; English.
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid, described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
SQ Sequence 1653 BP; 419 A; 430 C; 465 G; 339 T; 0 U; 0 Other;

Query Match 60.6%; Score 18.8; DB 5; Length 1653;
Best Local Similarity 76.7%; Pred. No. 2.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ACTGAAGCGGTGACGACATATCAGACGGCT 31
Db 1558 ACTGATCGGTGACTGCAGATGAGATCGCT 1587

RESULT 22
ID ACA01840
AC ACA01840 standard; DNA; 1653 BP.
XX ACA01840;
XX
DT 04-JUN-2003 (first entry)
XX
DE C. glutamicum derived ORF SEQ ID 1831.
XX
KW Coryneform; nucleic acid array; fermentation; culture; ds.
XX Corynebacterium glutamicum.
XX DE10128510-A1.
XX
PD 19-DEC-2002.
XX
PF 13-JUN-2001; 2001DE-01028510.
XX
PR 13-JUN-2001; 2001DE-01028510.

XX (DEGS) DEGUSSA AG.
XX Farwick M, Moeckel B, Pfefferle W, Bathe B, Huthmacher K;
XX WPI; 2003-279970/28.
XX New nucleic acid array useful for monitoring mRNA expression of
PT Corynebacterium glutamicum during fermentation, comprising nucleic acid
PT from Corynebacterium glutamicum.
XX Claim 1; Page 611-612; 709pp; German.
XX This invention describes a novel nucleic acid array involving
CC Corynebacterium glutamicum polynucleotides. The arrays are used to
CC analyse C. glutamicum, particularly for monitoring a fermentation process
CC to determine expression levels of C. glutamicum cellular mRNA. Such
CC monitoring particularly differentiates between expression levels of
CC different strains of C. glutamicum and allows the adjustment of different
CC culture and fermentation conditions. ACA00010-ACA02188 represent C.
CC glutamicum derived polynucleotides described in the disclosure of the
CC invention
XX
SQ Sequence 1653 BP; 419 A; 430 C; 465 G; 339 T; 0 U; 0 Other;

Query Match 60.6%; Score 18.8; DB 8; Length 1653;
Best Local Similarity 76.7%; Pred. No. 2.3e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACGACATATCAGACGGCT 31
Db 1558 ACTGATCGGTGACTGCAGATGAGATCGCT 1587

RESULT 23
ID AAK94137
AC AAK94137 standard; cDNA; 2324 BP.
XX
AC AAK94137;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human full-length cDNA, SEQ ID NO: 2642.
XX
KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS Homo sapiens.
XX
FN EP1130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-00114089.
XX
PR 08-JUL-1999; 99JP-00194486.
PR 11-JAN-2000; 2000JP-00118774.
PR 02-MAY-2000; 2000JP-00183765.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI; 2001-524255/58.
DR P-PSDB; AAM93227.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their use
PT in genetic manipulation.
XX
PS Claim 8; SEQ ID NO 2642; 1380pp + Sequence Listing; English.
XX
CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been isolated

CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
 CC been determined. Primers for synthesizing the full length cDNA are useful
 CC for clarifying the function of the protein encoded by the cDNA. The full
 CC length clones were obtained by construction of full length enriched cDNA
 CC libraries that were synthesized by the oligo-capping method. The primers
 CC enable the production of the full length cDNA easily without any special
 CC methods. The present sequence is a full length human cDNA of the
 CC invention. Note: the sequence data for this patent did not form part of
 CC the printed specification, but was obtained in CD-ROM format directly
 CC from EPO
 XX
 SQ Sequence 2324 BP; 493 A; 587 C; 657 G; 587 T; 0 U; 0 Other;

Query Match 60.6%; Score 18.8; DB 4; Length 2324;
 Best Local Similarity 76.7%; Pred. No. 2.5e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGCGGCT 31
 Db 911 ATTGAAGCTGGTGCACAGGATGTGAGCGGTT 940

RESULT 24
 ADL30609
 ID ADL30609 standard; cDNA; 2324 BP.

XX AC ADL30609;

XX 20-MAY-2004 (first entry)

XX DE Full length human cDNA clone SeqID 2642.

XX KW human; medicine; signal transduction; glycoprotein; transcription;
 KW oligo-capping method; ss; gene.

XX OS Homo sapiens.

XX PN EP1396543-A2.

XX PD 10-MAR-2004.

XX PP 07-JUL-2000; 2003EP-00025638.

XX PR 08-JUL-1999; 99JP-00194486.

XX PR 11-JAN-2000; 2000JP-00118774.

XX PR 02-MAY-2000; 2000JP-00183865.

XX PR 07-JUL-2000; 2000EP-00114089.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX PI Ota T, Nishikawa T, Isogai T, Hayaashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

XX DR WPI; 2004-204755/20.

XX DR P-PSDB; ADL30610.

XX PT New oligonucleotide primers (830 cDNAs) useful for synthesizing full
 PT length human cDNAs.

XX PS Example 1; SEQ ID NO 2642; 1340pp; English.

XX CC This invention relates to a novel primers useful for synthesizing full
 CC length cDNA molecules that encode human proteins. Specifically, it refers
 CC to secretory or membrane proteins that are potential therapeutic agents/
 CC target molecules in the field of medicine, and in particular genes
 CC encoding proteins that are associated with signal transduction,
 CC glycoproteins and transcription. The present invention describes a method
 CC for efficiently cloning a full length human cDNA from both the 5' and 3'
 CC ends using the oligo-capping method. This polynucleotide sequence is a
 CC full length human cDNA clone of the invention.

XX SQ Sequence 2324 BP; 493 A; 587 C; 657 G; 587 T; 0 U; 0 Other;

Query Match 60.6%; Score 18.8; DB 12; Length 2324;
 Best Local Similarity 76.7%; Pred. No. 2.5e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGCGGCT 31
 Db 911 ATTGAAGCTGGTGCACAGGATGTGAGCGGTT 940

RESULT 25

ABL61989

ID ABL61989 standard; DNA; 2640 BP.

XX AC ABL61989;

XX 15-MAY-2002 (first entry)

XX DE Colon adenocarcinoma related gene sequence SEQ ID NO:326.

XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.

XX OS Homo sapiens.

XX PN WO200194629-A2.

XX PD 13-DEC-2001.

XX PP 30-MAY-2001; 2001WO-US010838.

XX PR 05-JUN-2000; 2000US-0209473P.

XX PR 05-JUN-2000; 2000US-0209531P.

XX PR 18-SEP-2000; 2000US-0233133P.

XX PR 20-SEP-2000; 2000US-0233617P.

XX PR 20-SEP-2000; 2000US-0234009P.

XX PR 20-SEP-2000; 2000US-0234034P.

XX PR 22-SEP-2000; 2000US-0234509P.

XX PR 25-SEP-2000; 2000US-0234567P.

XX PR 25-SEP-2000; 2000US-0234923P.

XX PR 25-SEP-2000; 2000US-0235077P.

XX PR 25-SEP-2000; 2000US-0235082P.

XX PR 25-SEP-2000; 2000US-0235134P.

XX PR 25-SEP-2000; 2000US-0235280P.

XX PR 26-SEP-2000; 2000US-0235637P.

XX PR 26-SEP-2000; 2000US-0235638P.

XX PR 27-SEP-2000; 2000US-0235711P.

XX PR 27-SEP-2000; 2000US-0235720P.

XX PR 27-SEP-2000; 2000US-0235840P.

XX PR 27-SEP-2000; 2000US-0235863P.

XX PR 28-SEP-2000; 2000US-0236028P.

XX PR 28-SEP-2000; 2000US-0236032P.

XX PR 28-SEP-2000; 2000US-0236033P.

XX PR 28-SEP-2000; 2000US-0236034P.

XX PR 28-SEP-2000; 2000US-0236109P.

XX PR 29-SEP-2000; 2000US-0236111P.

XX PR 29-SEP-2000; 2000US-0236842P.

XX PR 29-SEP-2000; 2000US-0236891P.

XX PR 02-OCT-2000; 2000US-0237172P.

XX PR 02-OCT-2000; 2000US-0237173P.

XX PR 02-OCT-2000; 2000US-0237278P.

XX PR 02-OCT-2000; 2000US-0237294P.

XX PR 02-OCT-2000; 2000US-0237295P.

XX PR 03-OCT-2000; 2000US-0237316P.

XX PR 03-OCT-2000; 2000US-0237425P.

XX PR 03-OCT-2000; 2000US-0237598P.

XX PR 03-OCT-2000; 2000US-0237604P.

XX PR 03-OCT-2000; 2000US-0237606P.

XX PR 03-OCT-2000; 2000US-0237608P.

XX PR 01-NOV-2000; 2000US-0244867P.

PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
DR WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 326; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
CC tumour
XX
XX Sequence 2640 BP; 665 A; 647 C; 656 G; 672 T; 0 U; 0 Other;
SQ
Query Match 60.6%; Score 18.8; DB 6; Length 2640;
Best Local Similarity 76.7%; Pred. No. 2.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACAGCATATCAGACGGCT 31
DB 189 ACTGATCGCATGAGAGATATATACAGGGCT 218
RESULT 26
ADN03923
ID ADN03923 standard; cDNA; 2640 BP.
XX
AC ADN03923;
XX
DT 01-JUL-2004 (first entry)
XX
DE Antipsoriatic cDNA sequence #161.
XX
KW ds; gene; antipsoriatic; gene therapy; psoriasis; diagnosis.
XX
OS Homo sapiens.
XX
XX WO2004028479-A2.
XX
XX
PD 08-APR-2004.
XX
XX 25-SEP-2003; 2003WO-US030907.
XX
XX 25-SEP-2002; 2002US-0414006P.
XX
XX (GETH) GENENTECH INC.
XX
XX Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI Wu TD;
XX
XX WPI; 2004-305105/28.
DR
DR P-PSDB; ADN03924.

XX New PRO nucleic acid or polypeptide, useful for preparing a
PT pharmaceutical composition for diagnosing or treating psoriasis in a
PT mammal.
XX
XX Claim 1; SEQ ID NO 317; 3069pp; English.
XX
XX The invention relates to novel polynucleotide and polypeptides for
CC treating psoriasis or a sequence having at least 80% identity to the
CC above sequences. The nucleic acid is useful for preparing a composition
CC for diagnosing or treating psoriasis in a mammal. This sequence
CC corresponds to one of the polynucleotides of the invention.
XX
XX Sequence 2640 BP; 665 A; 647 C; 656 G; 672 T; 0 U; 0 Other;
SQ
Query Match 60.6%; Score 18.8; DB 12; Length 2640;
Best Local Similarity 76.7%; Pred. No. 2.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACAGCATATCAGACGGCT 31
DB 189 ACTGATCGCATGAGAGATATATACAGGGCT 218
RESULT 27
ADP10518
ID ADP10518 standard; DNA; 2640 BP.
XX
AC ADP10518;
XX
XX 12-AUG-2004 (first entry)
DT
XX
DE Reference mRNA sequences for marker probe #195.
XX
XX transplant rejection; immune system; rheumatoid arthritis; lupus;
KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS; ss.
XX
OS Homo sapiens.
XX
XX WO2004042346-A2.
XX
PD 21-MAY-2004.
XX
XX 24-APR-2003; 2003WO-US012946.
XX
XX 24-APR-2002; 2002US-00131831.
PR
XX 20-DEC-2002; 2002US-00325899.
XX
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.
XX
XX Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;
PI Rosenberg S;
XX
XX WPI; 2004-400724/37.
DR
XX
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant
PT rejection, in an individual, comprises detecting the expression level of
PT the genes.
XX
XX Claim 80; SEQ ID NO 527; 1762pp; English.
XX
XX The present invention relates to diagnosing or monitoring transplant
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual
CC comprises detecting the expression level of one or more genes. The
CC methods, system and kits are useful in diagnosing or monitoring
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic
CC islet, lung, bone marrow or stem cell transplant rejection,
CC xenotransplant rejection or mechanical organ replacement rejection, in an
CC individual. The method is also useful in assessing the immune status of
CC an individual. The methods are also useful in diagnosing and monitoring
CC diseases that involve the immune system, e.g. rheumatoid arthritis,
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or


```
RESULT 30
AAH64966
ID AAH64966 standard; DNA; 349980 BP.
XX
XX
AC AAH64966;
XX
DT 26-SEP-2001 (first entry)
XX
XX C glutamicum coding sequence fragment SEQ ID NO: 1.
XX
XX Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; ds.
XX
XX Corynebacterium glutamicum.
OS
XX
XX EF1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
PR 07-APR-2000; 2000JP-00159162.
PR 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
XX
XX Novel polynucleotides derived from Coryneform bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX Claim 7; SEQ ID NO 1; 246bp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
CC sequences from the Coryneform bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of coryneform bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Coryneform bacterium, and identifying a homologue of a gene derived from
CC Coryneform bacterium. Coryneform bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
XX SQ Sequence 349980 BP; 79703 A; 91547 C; 98381 G; 80349 T; 0 U; 0 Other;
Query Match 60.6%; Score 18.8; DB 5; Length 349980;
Best Local Similarity 76.7%; Pred. No. 6e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAACGGGTGACAGCATATCAGACGGCT 31
| | | | | | | | | | | | | | | | | | | | |
Db 62226 ACTGATGCGGTGACTGCAGATGAGATCGCT 62255
| | | | | | | | | | | | | | | | | | | | |
RESULT 31
AAS68064
ID AAS68064 standard; cDNA; 1398 BP.
XX
XX
AC AAS68064;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #3868.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW
```

```
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG03877.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 3868; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have application in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1398 BP; 424 A; 340 C; 479 G; 155 T; 0 U; 0 Other;
Query Match 60.0%; Score 18.6; DB 5; Length 1398;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 GAAGCGGTGACAGCATATCAGACGG 29
| | | | | | | | | | | | | | | | | | | | |
Db 1329 GCAGCGGAGACAGCAGCAGACGG 1353
| | | | | | | | | | | | | | | | | | | | |
RESULT 32
ABN68834
ID ABN68834 standard; DNA; 1464 BP.
XX
XX
AC ABN68834;
XX
XX 01-JUL-2002 (first entry)
XX
XX Streptococcus polynucleotide SEQ ID NO 5581.
DE
XX Streptococcus; GAS; GBS; Group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
XX antinflammatory; infection; vaccine; meningitis; gene therapy; ds.
```

```
XX OS Streptococcus pyogenes.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB004789.
XX PR 27-OCT-2000; 2000GB-00026333.
XX PR 24-NOV-2000; 2000GB-00028727.
XX PR 07-MAR-2001; 2001GB-00005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Maignani V, Margarit Y Rosl, Grandi G, Fraser C;
XX PI Tettelin H;
XX DR WPI; 2002-352536/38.
XX DR P-PSDB; ABP28203.
XX PT New Streptococcus protein for the treatment or prevention of infection or
XX PT disease caused by Streptococcus bacteria, such as meningitis, and for
XX PT detecting a compound that binds to the protein.
XX PS Claim 7; Page 3721; 4525pp; English.
XX CC The invention relates to a protein (ABP25413-ABP30895) from group B
XX CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX CC the specification. The proteins have antibacterial and antiinflammatory
XX CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
XX CC antibodies that bind (I) are used in the manufacture of medicaments for
XX CC the treatment or prevention of infection or disease caused by
XX CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX CC Nucleic acids encoding (I) are used to detect Streptococcus in a
XX CC biological sample. (I) is used to determine whether a compound binds to
XX CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
XX CC used as a vaccine or diagnostic composition. The disease caused by
XX CC Streptococcus that is prevented or treated may be meningitis. Nucleic
XX CC acid encoding (I) may be used to recombinantly produce (I) and may be
XX CC used in gene therapy. Antibodies to (I) are used for affinity
XX CC chromatography, immunoassays, and distinguishing/identifying
XX CC Streptococcus proteins
XX SQ Sequence 1464 BP; 408 A; 304 C; 327 G; 425 T; 0 U; 0 Other;
Query Match 60.0%; Score 18.6; DB 6; Length 1464;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 TGAAGCGGTGACGATATCAGCG 28
Db 1410 TGAAGCGGTGACGATATCACAAG 1434
RESULT 33
ACAS0663
ID ACAS0663 standard; DNA; 1467 BP.
XX AC ACAS0663;
XX DT 19-JUN-2003 (first entry)
XX DE Prokaryotic essential gene #32320.
XX KW Antisense; ds; prokaryotic essential gene; cell proliferation;
XX KW drug design; gene.
XX OS Streptococcus pyogenes.
XX PN WO200277183-A2.
```

```
XX PD 03-OCT-2002.
XX XX 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX PA (ELIT-) ELITRA PHARM INC.
XX XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen Kl, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX DR WPI; 2003-029926/02.
XX DR P-PSDB; ABU46793.
XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 14; SEQ ID NO 38533; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than S. aureus, S. typhimurium,
XX CC K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX CC prokaryotic essential genes. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 1467 BP; 410 A; 304 C; 327 G; 426 T; 0 U; 0 Other;
Query Match 60.0%; Score 18.6; DB 8; Length 1467;
Best Local Similarity 84.0%; Pred. No. 2.8e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 4 TGAAGCGGTGACGATATCAGCG 28
Db 1410 TGAAGCGGTGACGATATCACAAG 1434
RESULT 34
AAC09705/c
ID AAC09705 standard; cDNA; 251 BP.
XX XX AAC09705;
XX AC AAC09705;
```

XX 06-OCT-2000 (first entry)
DT XX
DE XX
DE Human secreted protein 5' EST, SEQ ID NO: 13780.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
KW
XX Homo sapiens.
OS
XX EP1033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-00200610.
PF
XX 26-FEB-1999; 99US-0122487P.
PR
XX (GEST) GENSET.
PA
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI; 2000-500381/45.
DR
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
PT
XX Claim 1; SEQ ID NO 13780; 71pp + Sequence Listing; English.
PS
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 251 BP; 67 A; 68 C; 32 G; 83 T; 0 U; 1 Other;

Query Match 59.4%; Score 18.4; DB 3; Length 251;
Best Local Similarity 78.6%; Pred. No. 2.5e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGCGG 29
DB 148 AGTGAAGTGGAGACACGGAATCAGGCTG 121

RESULT 35.
ABV38251
ID ABV38251 standard; cDNA; 438 BP.
XX
AC ABV38251;
XX
DT 16-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 38242.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200160860-A2.
PN
XX

PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JB;
PI WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
PT
XX Claim 1; Page 7803; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 438 BP; 96 A; 118 C; 130 G; 94 T; 0 U; 0 Other;

Query Match 59.4%; Score 18.4; DB 5; Length 438;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGCGG 30
DB 125 GTGAAGCGGTGACAGCATATCAGCGG 152

RESULT 36
ABV24200
ID ABV24200 standard; cDNA; 485 BP.
XX
AC ABV24200;
XX
DT 16-SEP-2002 (first entry)
DE Human prostate expression marker cDNA 24191.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
XX WO200160860-A2.
PN
XX 23-AUG-2001.
PD
XX 20-FEB-2001; 2001WO-US005171.
PF
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR

PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 4512; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 485 BP; 97 A; 139 C; 127 G; 121 T; 0 U; 1 Other;

Query Match 59.4%; Score 18.4; DB 5; Length 485;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
DB 76 GTGAGCAGTGACAGCATGCTGACGGC 103

RESULT 37
ADS56949
ID ADS56949 standard; cDNA; 1086 BP.
XX
AC ADS56949;
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polynucleotide #8936.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polynucleotide; gene; ss.
XX
OS Bacteria.
XX
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.
XX
PT New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 32623; 122pp; English.
XX
CC The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polynucleotide used in
CC the scope of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 1086 BP; 229 A; 329 C; 309 G; 219 T; 0 U; 0 Other;

Query Match 59.4%; Score 18.4; DB 13; Length 1086;
Best Local Similarity 78.6%; Pred. No. 3.3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
DB 269 GTGATGCGGTGACTGCGTATGAACAGC 296

RESULT 38
ABV25356
ID ABV25356 standard; cDNA; 1289 BP.
XX
AC ABV25356;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 25347.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
DR Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 5003; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1289 BP; 346 A; 300 C; 312 G; 331 T; 0 U; 0 Other;
Query Match 59.4%; Score 18.4; DB 5; Length 1289;
Best Local Similarity 78.6%; Pred. No. 3.4e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 GTGAAGCGGTGACGATATCAGACGGC 30
Db 753 GTGGAGCAGTGACGATGTCGACGGC 780
RESULT 39
ID ABV25037 standard; cDNA; 1289 BP.
XX
XX ABV25037;
XX
XX 16-SEP-2002 (first entry)
XX Human prostate expression marker cDNA 25028.
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 4854-4855; 11750pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1289 BP; 346 A; 300 C; 312 G; 331 T; 0 U; 0 Other;
Query Match 59.4%; Score 18.4; DB 5; Length 1289;
Best Local Similarity 78.6%; Pred. No. 3.4e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 GTGAAGCGGTGACGATATCAGACGGC 30
Db 753 GTGGAGCAGTGACGATGTCGACGGC 780
RESULT 40
AAD24919
ID AAD24919 standard; DNA; 1674 BP.
XX
XX AAD24919;
XX
XX 12-MAR-2002 (first entry)
XX Human T cell receptor gamma promoter DNA.
DE Human; growth inhibitory gene; retinoid; retinoic acid response element;
KW RARE site; therapy; promyelocytic leukaemia; cancer chemoprevention;
KW cytostatic; T cell receptor gamma promoter; ds.
XX Homo sapiens.
XX WO200192578-A2.
XX
XX 06-DEC-2001.
XX
XX 25-MAY-2001; 2001WO-US017161.
XX
XX 26-MAY-2000; 2000US-0207535P.
XX
XX (UNII) UNIV ILLINOIS FOUND.
XX
XX Roninson IB, Dokmanovic M, Chang B;
XX WPI; 2002-075474/10.
XX
XX Expression construct encoding cellular genes, under control of a promoter
PT regulated by retinoids and cells comprising the construct for identifying
PT compounds that induce expression of the genes useful in treating cancer.
XX
XX Claim 4; Page 50-51; 64pp; English.
PS
XX The patent discloses growth inhibitory genes induced by retinoids. The
CC invention also relates to recombinant expression constructs that express
CC a reporter gene under the transcriptional control of a promoter for a
CC gene which is expressed by retinoid induction. The promoter does not
CC contain a retinoic acid response elements (RARE) site. The invention
CC further relates to reagents and methods for identifying compounds other
CC than retinoids that modulate the expression of cellular genes. These
CC compounds are useful for treating cancers such as promyelocytic leukaemia
CC and cancer chemoprevention. The present sequence is human T cell receptor
CC gamma promoter DNA used in the invention

XX
SQ Sequence 1674 BP; 366 A; 511 C; 328 G; 469 T; 0 U; 0 Other;
Query Match 59.4%; Score 18.4; DB 6; Length 1674;
Best Local Similarity 78.6%; Pred. No. 3.5e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
DB 34 GTGGAGCAGTGACAGCATGTCTGCAGGC 61
Search completed: June 4, 2005, 07:29:05
Job time : 237.976 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:48:59 ; Search time 68.9704 Seconds
(without alignments)
735.454 Million cell updates/sec

Title: US-09-674-277-26
Perfect score: 31
Sequence: 1 tagtgaagcgggtgacagcatatcagcggct 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCUTS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	18.8	60.6	236	4	US-09-513-999C-21547 Sequence 21547, A
2	18.8	60.6	601	4	US-09-949-016-198524 Sequence 198524, A
3	18.8	60.6	1437	4	US-09-252-991A-8620 Sequence 8620, Ap
4	18.8	60.6	1506	4	US-09-252-991A-8602 Sequence 8602, Ap
5	18.8	60.6	1626	4	US-09-252-991A-8642 Sequence 8642, Ap
6	18.8	60.6	2716	4	US-09-949-016-5748 Sequence 5748, Ap
7	18.8	60.6	141115	4	US-09-949-016-17490 Sequence 17490, A
8	18.8	60.6	767677	4	US-09-949-016-12147 Sequence 12147, A
9	18.8	60.6	767677	4	US-09-949-016-17361 Sequence 17361, A
10	18.6	60.0	2572	3	US-09-221-017B-225 Sequence 225, App
11	18.4	59.4	251	4	US-09-513-999C-13780 Sequence 13780, A
12	18.4	59.4	1674	4	US-09-865-879-8 Sequence 8, Appli
13	18.4	59.4	67156	4	US-09-949-016-12284 Sequence 12284, A
14	18.4	59.4	67157	4	US-09-949-016-16558 Sequence 16558, A
15	18.4	59.4	78491	4	US-09-949-016-15132 Sequence 15132, A
16	18.2	58.7	1470	1	US-08-124-674-1 Sequence 1, Appli
17	18.2	58.7	1470	1	US-08-589-893-1 Sequence 1, Appli
18	18.2	58.7	1470	1	US-08-589-893-3 Sequence 3, Appli
19	18.2	58.7	1470	1	US-08-589-893-5 Sequence 5, Appli
20	18.2	58.7	1470	1	US-08-589-893-7 Sequence 7, Appli
21	18.2	58.7	1470	1	US-08-589-893-9 Sequence 9, Appli
22	18.2	58.7	1470	1	US-08-589-893-11 Sequence 11, Appli
23	18.2	58.7	1470	1	US-08-589-893-13 Sequence 13, Appli
24	18.2	58.7	1470	1	US-08-589-893-15 Sequence 15, Appli
25	18.2	58.7	1470	1	US-08-589-893-17 Sequence 17, Appli
26	18.2	58.7	1470	1	US-08-589-893-19 Sequence 19, Appli
27	18.2	58.7	1470	1	US-08-589-893-21 Sequence 21, Appli

C 28	18.2	58.7	1470	1	US-08-589-893-23	Sequence 23, Appli
C 29	18.2	58.7	1470	2	US-09-020-991-1	Sequence 1, Appli
C 30	18.2	58.7	1470	2	US-09-020-991-3	Sequence 3, Appli
C 31	18.2	58.7	1470	2	US-09-020-991-5	Sequence 5, Appli
C 32	18.2	58.7	1470	2	US-09-020-991-7	Sequence 7, Appli
C 33	18.2	58.7	1470	2	US-09-020-991-9	Sequence 9, Appli
C 34	18.2	58.7	1470	2	US-09-020-991-11	Sequence 11, Appli
C 35	18.2	58.7	1470	2	US-09-020-991-13	Sequence 13, Appli
C 36	18.2	58.7	1470	2	US-09-020-991-15	Sequence 15, Appli
C 37	18.2	58.7	1470	2	US-09-020-991-17	Sequence 17, Appli
C 38	18.2	58.7	1470	2	US-09-020-991-19	Sequence 19, Appli
C 39	18.2	58.7	1470	2	US-09-020-991-21	Sequence 21, Appli
C 40	18.2	58.7	1470	2	US-09-020-991-23	Sequence 23, Appli
C 41	18.2	58.7	1470	2	US-09-062-890-1	Sequence 1, Appli
C 42	18.2	58.7	1470	2	US-09-062-890-3	Sequence 3, Appli
C 43	18.2	58.7	1470	2	US-09-062-890-5	Sequence 5, Appli
C 44	18.2	58.7	1470	2	US-09-062-890-7	Sequence 7, Appli
C 45	18.2	58.7	1470	2	US-09-062-890-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-513-999C-21547
; Sequence 21547, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 21547
; LENGTH: 236
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-513-999C-21547

Query Match 60.6%; Score 18.8; DB 4; Length 236;
Best Local Similarity 76.7%; Pred. No. 29;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ACTGAAGCGGTGACAGCATATCAGACGGCT 31
||||| ||||| ||||| ||||| ||||| |||||
DB 109 ACTGATCGCATGAAAGATATATACAGGGCT 138

RESULT 2

US-09-949-016-198524/c
; Sequence 198524, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012


```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17490
; LENGTH: 14115
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(14115)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17490

Query Match      60.6%; Score 18.8; DB 4; Length 14115;
Best Local Similarity 76.7%; Pred. No. 1.4e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      2 AGTGAGCGGTGACGACATATCAGCGGCT 31
Db      6568 AGTGATCGCATGAAGAATATACAGCGGCT 6597

RESULT 8
US-09-949-016-12147
; Sequence 12147, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12147
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147

Query Match      60.6%; Score 18.8; DB 4; Length 767677;
Best Local Similarity 90.9%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      8 GCGGTGACGACATATCAGACGG 29
Db      454532 GCTGTGACAGCATAGCAGCGG 454553

RESULT 9
US-09-949-016-17361
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17361
; LENGTH: 767677
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(767677)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17361

Query Match      60.6%; Score 18.8; DB 4; Length 767677;
Best Local Similarity 90.9%; Pred. No. 2.1e+02;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      8 GCGGTGACGACATATCAGACGG 29
Db      454532 GCTGTGACAGCATAGCAGCGG 454553

RESULT 10
US-09-221-017B-225/c
; Sequence 225, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: ROSS, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
```

; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2572 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORYPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2572
US-09-221-017B-225

Query Match 60.0%; Score 18.6; DB 3; Length 2572;
Best Local Similarity 84.0%; Pred. No. 66;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 AGCGGTGACGACATATCAGCGGCT 31
Db 817 AGCGGTGACGACGACGACGCGGCT 793

RESULT 11
US-09-513-999C-13780/c
; Sequence 13780, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 13780
; LENGTH: 251
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 161
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-13780

Query Match 59.4%; Score 18.4; DB 4; Length 251;
Best Local Similarity 78.6%; Pred. No. 45;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACGACATATCAGCGG 29
Db 148 AGTGAAGTGGACGACGACGACGCGT 121

RESULT 12
US-09-865-879-8
; Sequence 8, Application US/09865879

; Patent No. 6767705
; GENERAL INFORMATION:
; APPLICANT: Roninson, Igor
; APPLICANT: Dokmanovic, Milos
; APPLICANT: Chang, Bey-Dih
; TITLE OF INVENTION: REAGENTS AND METHODS FOR IDENTIFYING AND MODULATING EXPRESSION OF
; TITLE OF INVENTION: REGULATED BY RETINOIDS
; FILE REFERENCE: 99,216-H
; CURRENT APPLICATION NUMBER: US/09/865,879
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,535
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: T cell receptor gamma promoter NCBI acc. number AC006033.2
US-09-865-879-8

Query Match 59.4%; Score 18.4; DB 4; Length 1674;
Best Local Similarity 78.6%; Pred. No. 73;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGACGGC 30
Db 34 GTGAAGCGGTGACGACATGTCTGCAGGC 61

RESULT 13
US-09-949-016-12284
; Sequence 12284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12284
; LENGTH: 67156
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12284

Query Match 59.4%; Score 18.4; DB 4; Length 67156;
Best Local Similarity 78.6%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGACGGC 30
Db 39734 GTGAAGGGGCGCCAGCACATCAGTCG 39761

RESULT 14
US-09-949-016-16558
; Sequence 16558, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16558
; LENGTH: 67157
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16558

Query Match          59.4%; Score 18.4; DB 4; Length 67157;
Best Local Similarity 78.6%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGCGC 30
Db 39734 GTGAAGGGAGCGCAGCATCATCATGGC 39761

RESULT 15
US-09-949-016-15132/c
; Sequence 15132, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15132
; LENGTH: 78491
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..-(78491)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15132

Query Match          59.4%; Score 18.4; DB 4; Length 78491;
Best Local Similarity 78.6%; Pred. No. 1.9e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACGACATATCAGCGG 29
Db 3514 AGTGCAGTGGTGACAGCATAGTCACTG 3487

RESULT 16
US-08-124-674-1/c
; Sequence 1, Application US/08124674
; Patent No. 5468632
; GENERAL INFORMATION:
; APPLICANT: Cantwell et al.
; TITLE OF INVENTION: Recombinant DNA Compounds and Expression Vectors Encoding para
; NUMBER OF SEQUENCES: 7
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: IN
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/124,674
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/975,206
; FILING DATE:
; APPLICATION NUMBER: 07/811,096
; FILING DATE: December 20, 1991
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-124-674-1

Query Match          58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
Db 210 CAGTCAGAGCAATCAGACGGCT 188

RESULT 17
US-08-589-893-1/c
; Sequence 1, Application US/08589893
; Patent No. 5741691
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,893
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 1:
```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 0-Wtpnb
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
;
US-08-589-893-1
Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
| |||| |||| |||| |||| ||||
Db 210 CAGTGAGACCAATCAGACGGCT 188

RESULT 18
US-08-589-893-3/c
; Sequence 3, Application US/08589893
; Patent No. 5741691
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poma, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,893
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 1-lb9
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
;

Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
| |||| |||| |||| |||| ||||
Db 210 CAGTGAGACCAATCAGACGGCT 188

RESULT 19
US-08-589-893-5/c
; Sequence 5, Application US/08589893
; Patent No. 5741691
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poma, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,893
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 2-19E10
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
;

Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
| |||| |||| |||| |||| ||||
Db 210 CAGTGAGACCAATCAGACGGCT 188

RESULT 20
US-08-589-893-7/c
; Sequence 7, Application US/08589893
```

```

US-08-589-893-3
Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
| |||| |||| |||| |||| ||||
Db 210 CAGTGAGACCAATCAGACGGCT 188

RESULT 19
US-08-589-893-5/c
; Sequence 5, Application US/08589893
; Patent No. 5741691
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poma, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,893
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 2-19E10
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
;

Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
| |||| |||| |||| |||| ||||
Db 210 CAGTGAGACCAATCAGACGGCT 188

RESULT 20
US-08-589-893-7/c
; Sequence 7, Application US/08589893
```



```

; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 4-43e7
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-08-589-893-11

```

Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0

Qy 9 CGGTGACAGCATATCAGACGGCT 31
 | | | | |
Db 210 CAGTGAGAGCAATCAGACGGCT 188

```

RESULT 23
US-08-589-893-13/c
; Sequence 13, Application US/08589893
; Patent No. 5741691
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZENE
; TITLE OF INVENTION: ACTIVITY IN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pons, Smith, Lande
; STREET: 2029 Century Park East
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 107260
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1299

/ 12574X: 13107 2177-12574
 / INFORMATION FOR SEQ ID NO: 13:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 1470 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: DNA
 / HYPOTHEetical: NO
 / ANTI-SENSE: NO
 / ORIGINAL SOURCE:
 / ORGANISM:
 / INDIVIDUAL ISOLATE: 4-54b9

INDIVIDUAL ISOLATE: 4-54b9

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-08-589-893-13

```

Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0

QY 9 CGGTGACAGCATATCAGACGGCT 31
 | | | | |
Db 210 CAGTGAGAGCAATCAGACGGCT 188

```

RESULT 24
US-08-589-893-15/c
; Sequence 15, Application US/08589893
; Patent No. 5741691
;
; GENERAL INFORMATION:
;
; APPLICANT: Artold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZENE
; TITLE OF INVENTION: ACTIVITY IN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poma Smith, Lande
; STREET: 2029 Century Park East
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
;

```

```

?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatenIn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/589,893
?
?
```

/ FILING DATE: /
 / CLASSIFICATION: 435 /
 / ATTORNEY/AGENT INFORMATION: /
 / NAME: Oldenkamp, David J. /
 / REGISTRATION NUMBER: 29,421 /
 / REFERENCE/DOCKET NUMBER: 107260 /
 / TELECOMMUNICATION INFORMATION: /
 / TELEPHONE: (310) 788-5000 /
 / TELEFAX: (310) 277-1297 /
 /

```

/ SELEXPR: 13107 277 1257
/ INFORMATION FOR SEQ ID NO: 15:
/   SEQUENCE CHARACTERISTICS:
/     LENGTH: 1470 base pairs
/     TYPE: nucleic acid
/     STRANDEDNESS: single
/     TOPOLOGY: linear
/     MOLECULE TYPE: DNA
/     HYPOTHEetical: NO
/     ANTI-SENSE: NO
/     ORIGINAL SOURCE:
/

```

```

; ORGANISM:
; INDIVIDUAL ISOLATE: 2-13f3
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-08-589-893-15

```

Query Match 58.7%; Score 18.2; DB 1; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0

QY
9 CCGTGACAGCATATCAGACGGCT 31

Dp
210 CAGTGAGAGCAATCAGACGGCT 188

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0;

QY 9 CGGTGACAGCATATCAGACGGCT 31
 | | | | | | | | | |
Dd 210 CAGTGAGAGCAAAATCAGACGGCT 188

RESULT 30
US-09-020-991-3/c
; Sequence 3, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poma, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20: Conservative 0; Mismatches 3; Indels 0

QY 9 CGGTGACAGCATATCAGACGGCT 31
| | | | | | | | | |
Dp 210 CAGTGAGAGCAATCAGACGGCT 188

RESULT 31
US-09-020-991-5/c
; Sequence 5, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:

APPLICANT: Arnold, Frances H.
APPLICANT: Moors, Jeffrey C.
TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Poms, Smith, Lande & Rose
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020, 991
FILING DATE:

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20: Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACAGCATATCAGACGGCT 31
 | | | | | | | | | |
Dh 210 CACTGAGAGCAATCAGACGGCT 188

RESULT 32
US-09-020-991-7/c
; Sequence 7, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA

ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,991
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/589,893
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 107260
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1470 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM:
INDIVIDUAL ISOLATE: 3-10c4
FEATURE:
NAME/KEY: CDS
LOCATION: complement (1..1470)
US-09-020-991-7

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
| | | | | | | | | | | | | | | | | | | | | |
Db 210 CAGTGAGAGCAATCAGACGGCT 188

RESULT 33
US-09-020-991-9/c
Sequence 9, Application US/09020991
Patent No. 5906930
GENERAL INFORMATION:
APPLICANT: Arnold, Frances H.
TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Poms, Smith, Lande & Rose
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,991
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/589,893
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 107260
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1470 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM:
INDIVIDUAL ISOLATE: 4-38b9
FEATURE:
NAME/KEY: CDS
LOCATION: complement (1..1470)
US-09-020-991-9

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGACGGCT 31
| | | | | | | | | | | | | | | | | | | | | |
Db 210 CAGTGAGAGCAATCAGACGGCT 188

RESULT 34
US-09-020-991-11/c
Sequence 11, Application US/09020991
Patent No. 5906930
GENERAL INFORMATION:
APPLICANT: Arnold, Frances H.
TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Poms, Smith, Lande & Rose
STREET: 2029 Century Park East, Suite 3800
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,991
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/589,893
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Oldenkamp, David J.
REGISTRATION NUMBER: 29,421
REFERENCE/DOCKET NUMBER: 107260
TELECOMMUNICATION INFORMATION:
TELEPHONE: (310) 788-5000
TELEFAX: (310) 277-1297
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 4-43e7
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-09-020-991-11

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGCGGCT 31
DB 210 CAGTGAGAGCAATCAGCGGCT 188

RESULT 35
US-09-020-991-13/c
; Sequence 13, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,991
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/589,893
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 4-54b9
; FEATURE:

; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-09-020-991-13

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGCGGCT 31
DB 210 CAGTGAGAGCAATCAGCGGCT 188

RESULT 36
US-09-020-991-15/c
; Sequence 15, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,991
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/589,893
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 2-13f3
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-09-020-991-15

Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 CGGTGACGACATATCAGCGGCT 31
DB 210 CAGTGAGAGCAATCAGCGGCT 188

RESULT 37
US-09-020-991-17/c
; Sequence 17, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,991
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/589,893
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 2-23e1
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-09-020-991-17
Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 9 CGGTGACGACATATCAGCGGCT 31
Db 210 CAGTGAGAGCAAAATCAGACGGCT 188
RESULT 38
US-09-020-991-19/c
; Sequence 19, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA

; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,991
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/589,893
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.
; REGISTRATION NUMBER: 29,421
; REFERENCE/DOCKET NUMBER: 107260
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (310) 788-5000
; TELEFAX: (310) 277-1297
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1470 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM:
; INDIVIDUAL ISOLATE: 4-53d5
; FEATURE:
; NAME/KEY: CDS
; LOCATION: complement (1..1470)
US-09-020-991-19
Query Match 58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 9 CGGTGACGACATATCAGCGGCT 31
Db 210 CAGTGAGAGCAAAATCAGACGGCT 188
RESULT 39
US-09-020-991-21/c
; Sequence 21, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/09/020,991
, FILING DATE:
, CLASSIFICATION:
, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/589,893
, FILING DATE:
, ATTORNEY/AGENT INFORMATION:
, NAME: Oldenkamp, David J.
, REGISTRATION NUMBER: 29,421
, REFERENCE/DOCKET NUMBER: 107260
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (310) 788-5000
, TELEFAX: (310) 277-1297
, INFORMATION FOR SEQ ID NO: 21:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 1470 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: single
, TOPOLOGY: linear
, MOLECULE TYPE: DNA
, HYPOTHETICAL: NO
, ANTI-SENSE: NO
, ORIGINAL SOURCE:
, ORGANISM:
, INDIVIDUAL ISOLATE: 5-1a12
, FEATURE:
, NAME/KEY: CDS
, LOCATION: complement (1..1470)
US-09-020-991-21

```

```
Query Match      58.7%; Score 18.2; DB 2; Length 1470;
Best Local Similarity 87.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

Qy 9 CGGTGACAGCATATCAGACGGCT 31
| | | | | | | | | |
Db 210 CAGTGAGAGCAATCAGACGGCT 188

RESULT 40
US-09-020-991-23/c
; Sequence 23, Application US/09020991
; Patent No. 5906930
; GENERAL INFORMATION:
; APPLICANT: Arnold, Frances H.
; APPLICANT: Moore, Jeffrey C.
; TITLE OF INVENTION: PARA-NITROBENZYL ESTERASES WITH ENHANCED
; TITLE OF INVENTION: ACTIVITY IN AQUEOUS AND NONAQUEOUS MEDIA
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Poms, Smith, Lande & Rose
; STREET: 2029 Century Park East, Suite 3800
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/020,991
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/589,893
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Oldenkamp, David J.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 07:08:50 ; Search time 295.509 Seconds
(without alignments)
644.888 Million cell updates/sec

Title: US-09-674-277-26

Perfect score: 31
Sequence: 1 tsgtgaagcggtagacagcatatcagacggt 31

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	19.4	62.6	642	13	US-10-027-632-230031 Sequence 230031,
C 2	19.4	62.6	642	17	US-10-027-632-230031 Sequence 230031,
C 3	19.4	62.6	650	13	US-10-027-632-227902 Sequence 227902,
C 4	19.4	62.6	650	17	US-10-027-632-227902 Sequence 227902,
C 5	19.4	62.6	1470	17	US-10-369-493-46995 Sequence 46995, A
6	19.4	62.6	2367	17	US-10-267-503-81 Sequence 81, Appl
7	19	61.3	2367	18	US-10-437-963-80893 Sequence 80893, A
8	19	61.3	3694	9	US-09-764-864-329 Sequence 329, App
9	19	61.3	71678	13	US-10-087-192-655 Sequence 655, App
10	18.8	60.6	624	13	US-10-027-632-113097 Sequence 113097,
11	18.8	60.6	624	17	US-10-027-632-113097 Sequence 113097,

12	18.8	60.6	653	13	US-10-027-632-32398 Sequence 32398, A
13	18.8	60.6	653	17	US-10-027-632-32398 Sequence 32398, A
14	18.8	60.6	1465	18	US-10-416-330-8 Sequence 8, Appl
15	18.8	60.6	1653	9	US-09-738-626-67 Sequence 67, Appl
16	18.8	60.6	2640	10	US-09-873-367C-326 Sequence 326, App
17	18.8	60.6	2640	19	US-10-843-641A-326 Sequence 326, App
18	18.8	60.6	3309400	9	US-09-738-626-1 Sequence 1, Appl
19	18.6	60.0	223	18	US-10-425-115-146059 Sequence 146059, A
20	18.6	60.0	1467	17	US-10-282-122A-38533 Sequence 38533, A
21	18.6	60.0	2572	13	US-10-194-163-225 Sequence 225, App
22	18.4	59.4	438	18	US-10-357-930-38270, A
23	18.4	59.4	485	18	US-10-357-930-38270, A
24	18.4	59.4	878	13	US-10-027-632-29410, A
25	18.4	59.4	878	17	US-10-027-632-29410, A
26	18.4	59.4	1086	17	US-10-369-493-32623 Sequence 32623, A
27	18.4	59.4	1289	18	US-10-357-930-25026 Sequence 25026, A
28	18.4	59.4	1289	18	US-10-357-930-25026 Sequence 25026, A
29	18.4	59.4	1638	13	US-10-027-632-97723 Sequence 97723, A
30	18.4	59.4	1638	17	US-10-027-632-97723 Sequence 97723, A
31	18.4	59.4	1674	10	US-09-865-879-8 Sequence 8, Appl
32	18.4	59.4	1799	15	US-10-101-510-492 Sequence 492, App
33	18.4	59.4	2291	17	US-10-412-699B-403 Sequence 403, App
34	18.4	59.4	3056	17	US-10-108-260A-1214 Sequence 1214, Ap
35	18.4	59.4	3826	15	US-10-128-714-223 Sequence 223, App
36	18.4	59.4	4473	18	US-10-437-963-10983 Sequence 10983, A
37	18.4	59.4	5394	15	US-10-278-536-223 Sequence 223, App
38	18.4	59.4	11122	11	US-09-984-429-394 Sequence 394, App
39	18.4	59.4	11122	11	US-09-984-429-444 Sequence 444, App
40	18.4	59.4	11307	9	US-09-070-927A-49 Sequence 49, Appl
41	18.4	59.4	230101	18	US-10-719-993-6829 Sequence 6829, Ap
42	18.4	59.4	9025608	15	US-10-156-761-1 Sequence 73604, A
43	18.2	58.7	561	17	US-10-424-599-73604 Sequence 73604, A
44	18.2	58.7	583	13	US-10-027-632-274155 Sequence 274155, A
45	18.2	58.7	583	17	US-10-027-632-274155 Sequence 274155, A

ALIGNMENTS

RESULT 1
US-10-027-632-230031/c
Sequence 230031, Application US/10027632
Publication No. US20020198371A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 230031
LENGTH: 642
TYPE: DNA
ORGANISM: Human
US-10-027-632-230031

Query Match 62.6%; Score 19.4; DB 13; Length 642;


```
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46995
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-10-369-493-46995

Query Match          62.6%; Score 19.4; DB 17; Length 1470;
Best Local Similarity 95.2%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 GTGACAGCATATCAGACGGCT 31
Db 208 GTGACAGCAATCAGACGGCT 188

RESULT 6
US-10-267-502-81
; Sequence 81, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 81
; LENGTH: 2367
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-267-502-81

Query Match          62.6%; Score 19.4; DB 17; Length 2367;
Best Local Similarity 79.3%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGGC 30
Db 1315 AGTGAAGCGGTGACAGCATTTGCGCGGC 1343

RESULT 7
US-10-437-963-80893
; Sequence 80893, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 80893
; LENGTH: 2367
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80473C.1
US-10-437-963-80893

Query Match          61.3%; Score 19; DB 18; Length 2367;
Best Local Similarity 81.5%; Pred. No. 1.7e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 5 GAAGCGGTGACAGCATATCAGACGGCT 31
Db 815 GAAGCGGTAACAGCGGTGCTGATGGCT 841

RESULT 8
US-09-764-864-329
; Sequence 329, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 329
; LENGTH: 3694
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-329

Query Match          61.3%; Score 19; DB 9; Length 3694;
Best Local Similarity 75.9%; Pred. No. 1.8e+02;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGGC 30
Db 1613 AGTGAAGCTGGAASCGCATATCAGAGTGC 1641

RESULT 9
US-10-087-192-655
; Sequence 655, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 655
; LENGTH: 71678
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(71678)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-655

Query Match          61.3%; Score 19; DB 13; Length 71678;
Best Local Similarity 81.5%; Pred. No. 2.4e+02;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGC 28
```

```
; SEQ ID NO 113097
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113097

Query Match      60.6%; Score 18.8; DB 17; Length 624;
Best Local Similarity 76.7%; Pred. No. 1.8e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 10
US-10-027-632-113097
; Sequence 113097, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113097
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113097

Query Match      60.6%; Score 18.8; DB 13; Length 624;
Best Local Similarity 76.7%; Pred. No. 1.8e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 11
US-10-027-632-113097
; Sequence 113097, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113097
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113097

Query Match      60.6%; Score 18.8; DB 13; Length 624;
Best Local Similarity 76.7%; Pred. No. 1.8e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 12
US-10-027-632-32398
; Sequence 32398, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32398
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-32398

Query Match      60.6%; Score 18.8; DB 13; Length 653;
Best Local Similarity 76.7%; Pred. No. 1.8e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

RESULT 13
US-10-027-632-32398
; Sequence 32398, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32398
; LENGTH: 653
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-32398

Query Match 60.6%; Score 18.8; DB 17; Length 653;
Best Local Similarity 76.7%; Pred. No. 1.8e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACGACATATCAGACGGCT 31
Db 270 AGTGAAGCGGTGACGACATATCAGACGGCT 299

RESULT 14
US-10-416-330-8
; Sequence 8, Application US/10416330
; Publication No. US20040197782A1
; GENERAL INFORMATION:
; APPLICANT: EICHMULLER, STEFAN
; APPLICANT: SCHADENDORF, DIRK
; APPLICANT: USENER, DIRK
; TITLE OF INVENTION: NOVEL MARKER FOR THE DIAGNOSIS AND THERAPY OF TUMORS
; FILE REFERENCE: 38485-0014
; CURRENT APPLICATION NUMBER: US/10/416,330
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: PCT/DE01/04229
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: DE 10055285.4
; PRIOR FILING DATE: 2000-11-08
; NUMBER OF SEQ ID NOS: 93
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 8
; LENGTH: 1465
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-416-330-8

Query Match 60.6%; Score 18.8; DB 18; Length 1465;
Best Local Similarity 76.7%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACGACATATCAGACGGCT 31
Db 493 AGTGAAGCGGTGACGACATATCAGACGGCT 522

RESULT 15
US-09-738-626-67
; Sequence 67, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO

; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 67
; LENGTH: 1653
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-67

Query Match 60.6%; Score 18.8; DB 9; Length 1653;
Best Local Similarity 76.7%; Pred. No. 2e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACGACATATCAGACGGCT 31
Db 1558 ACTGATCGCGTGACTGCAGATGAGATCGCT 1587

RESULT 16
US-09-873-367C-326
; Sequence 326, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Ebner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
; CURRENT APPLICATION NUMBER: US/09/873,367C
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,891
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 326
; LENGTH: 2640
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-367C-326

Query Match 60.6%; Score 18.8; DB 10; Length 2640;
Best Local Similarity 76.7%; Pred. No. 2.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 AGTGAAGCGGTGACGACATATCAGACGGCT 31
Db 189 AGTGAAGCGGTGACGACATATCAGACGGCT 218

RESULT 17
US-10-843-641A-326
; Sequence 326, Application US/10843641A
; Publication No. US20050064454A1

GENERAL INFORMATION:
APPLICANT: Avalon Pharmaceuticals, Inc.
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
TITLE OF INVENTION: Signature Gene Sets
FILE REFERENCE: 689290-189
CURRENT APPLICATION NUMBER: US/10/843,641A
CURRENT FILING DATE: 2004-05-12
PRIOR APPLICATION NUMBER: US/09/873,367
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US/09/954,531
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/09/954,456
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,436
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/962,832
PRIOR FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/09/964,824
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US/09/967,768
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US/09/968,007
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,347
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/09/969,708
PRIOR FILING DATE: 2001-10-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 8447
SOFTWARE: PatentIn version 3.0
SEQ ID NO 326
LENGTH: 2640
TYPE: DNA
ORGANISM: Homo sapiens
US-10-843-641A-326

Query Match 60.6%; Score 18.8; DB 19; Length 2640;
Best Local Similarity 76.7%; Pred. No. 2.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGGCT 31
|||||
Db 189 AGTGATGCGATGAAGAATATACAGGGCT 218
|||||

RESULT 18
US-09-738-626-1
Sequence 1, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENOH, AKIHO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 1

LENGTH: 3309400
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
US-09-738-626-1
Query Match 60.6%; Score 18.8; DB 9; Length 3309400;
Best Local Similarity 76.7%; Pred. No. 4e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2 ACTGAAGCGGTGACAGCATATCAGACGGCT 31
|||||
Db 62226 ACTGATGCGGTGACTGCAGATGAGATCGCT 62255
|||||
RESULT 19
US-10-425-115-146059/c
Sequence 146059, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 146059
LENGTH: 223
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_64720C.1
US-10-425-115-146059

Query Match 60.0%; Score 18.6; DB 18; Length 223;
Best Local Similarity 84.0%; Pred. No. 2e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 AGCGGTGACAGCATATCAGACGGCT 31
|||||
Db 205 AGCGGTGCGCAGCATGACACGGGT 181
|||||

RESULT 20
US-10-282-122A-38533
Sequence 38533, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26

```
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38533
; LENGTH: 1467
; TYPE: DNA
; ORGANISM: Streptococcus pyogenes
US-10-282-122A-38533

Query Match      60.0%; Score 18.6; DB 17; Length 1467;
Best Local Similarity 84.0%; Pred. No. 2.4e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      4 TGAAGCGGTGACAGCATATCAGACG 28
      |||||
Db      1410 TGAAGCGGTGACAGCATATCACAAG 1434

RESULT 21
US-10-194-163-225/c
; Sequence 225, Application US/10194163
; Publication No. US20020172976A1
; GENERAL INFORMATION:
; APPLICANT: ROSS, Bruce Carter
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYNUCLEOTIDES
; AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/194,163
; FILING DATE: 04-Nov-2002
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Basu, Shantanu
; REGISTRATION NUMBER: 43,318
; REFERENCE/DOCKET NUMBER: 529282000101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5995
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 225
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2572 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
```

```
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...2572
; SEQUENCE DESCRIPTION: SEQ ID NO: 225
US-10-194-163-225

Query Match      60.0%; Score 18.6; DB 13; Length 2572;
Best Local Similarity 84.0%; Pred. No. 2.6e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      7 AGCGGTGACAGCATATCAGACGGCT 31
      |||||
Db      817 AGCGGTGACAGCAGACGGCGGCT 793

RESULT 22
US-10-357-930-38270
; Sequence 38270, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38270
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-38270

Query Match      59.4%; Score 18.4; DB 18; Length 438;
Best Local Similarity 78.6%; Pred. No. 2.6e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY      3 GTGAGCGGTGACAGCATATCAGACGGC 30
      |||||
Db      125 GTGAGCAGTGACAGCATGTCTGCAGGC 152

RESULT 23
US-10-357-930-24189
; Sequence 24189, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
```

```
/ TITLE OF INVENTION: HUMAN PROSTATE CANCER
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ PRIOR FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 24189
/ LENGTH: 485
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 392
/ OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24189
```

```
Query Match 59.4%; Score 18.4; DB 18; Length 485;
Best Local Similarity 78.6%; Pred. No. 2.7e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
Db 76 GTGAGCAGTGTGACAGCATGTCTGCAGGC 103
```

```
RESULT 24
US-10-027-632-29410/c
/ Sequence 29410, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 29410
/ LENGTH: 878
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-29410
```

```
Query Match 59.4%; Score 18.4; DB 13; Length 878;
```

```
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
Db 151 GTGAAGGGGAGCCAGCATGTCTCATGGC 124
```

```
RESULT 25
US-10-027-632-29410/c
/ Sequence 29410, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 29410
/ LENGTH: 878
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-29410
```

```
Query Match 59.4%; Score 18.4; DB 17; Length 878;
Best Local Similarity 78.6%; Pred. No. 2.8e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
Db 151 GTGAAGGGGAGCCAGCATGTCTCATGGC 124
```

```
RESULT 26
US-10-369-493-32623
/ Sequence 32623, Application US/10369493
/ Publication No. US20030233675A1
/ GENERAL INFORMATION:
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Hinkle, Gregory J.
/ APPLICANT: Slater, Steven C.
/ APPLICANT: Goldman, Barry S.
/ APPLICANT: Chen, Xianfeng
/ TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
/ FILE REFERENCE: 38-10(52052)B
/ CURRENT APPLICATION NUMBER: US/10/369,493
/ CURRENT FILING DATE: 2003-02-28
/ PRIOR APPLICATION NUMBER: US 60/360,039
/ PRIOR FILING DATE: 2002-02-21
/ NUMBER OF SEQ ID NOS: 47374
/ SEQ ID NO 32623
/ LENGTH: 1086
/ TYPE: DNA
/ ORGANISM: Chloroflexus aurantiacus
US-10-369-493-32623
```



```
Query Match          59.4%; Score 18.4; DB 17; Length 1086;
Best Local Similarity 78.6%; Pred. No. 2.9e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGACGGC 30
Db 269 GTGATCGGTGACTGCGTATGAACAGC 296

RESULT 27
US-10-357-930-25026
; Sequence 25026, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25026
; LENGTH: 1289
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-25026

Query Match          59.4%; Score 18.4; DB 18; Length 1289;
Best Local Similarity 78.6%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGACGGC 30
Db 753 GTGGAGCAGTGACGACATGCTGCAGGC 780

RESULT 28
US-10-357-930-25345
; Sequence 25345, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
```

```
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25345
; LENGTH: 1289
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-25345

Query Match          59.4%; Score 18.4; DB 18; Length 1289;
Best Local Similarity 78.6%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGACGGC 30
Db 753 GTGGAGCAGTGACGACATGCTGCAGGC 780

RESULT 29
US-10-027-632-97723
; Sequence 97723, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 97723
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97723

Query Match          59.4%; Score 18.4; DB 13; Length 1638;
Best Local Similarity 78.6%; Pred. No. 3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGACGGC 30
Db 1488 GTGAAGGGGAGCGCATGTCATGCGC 1515

RESULT 30
US-10-027-632-97723
; Sequence 97723, Application US/10027632
; Publication No. US20030204075A9
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97723
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-97723

Query Match          59.4%; Score 18.4; DB 17; Length 1638;
Best Local Similarity 78.6%; Pred. No. 3e+02; Mismatches 0; Gaps 0;
Matches 22; Conservative 0; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
Db 1488 GTGAAGGGAGCGACAGCATGTCACATGGC 1515

RESULT 31
US-09-865-879-8
; Sequence 8, Application US/09865879
; Publication No. US20030180707A1
; GENERAL INFORMATION:
; APPLICANT: Roninson, Igor
; APPLICANT: Dokmanovic, Milos
; APPLICANT: Chang, Bey-Dih
; TITLE OF INVENTION: REAGENTS AND METHODS FOR IDENTIFYING AND MODULATING EXPRESSION OF
; TITLE OF INVENTION: REGULATED BY RETINOIDS
; FILE REFERENCE: 99,216-H
; CURRENT APPLICATION NUMBER: US/09/865,879
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,535
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 1674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: T cell receptor gamma promoter NCBI acc. number AC006033.2
US-09-865-879-8

Query Match          59.4%; Score 18.4; DB 10; Length 1674;
Best Local Similarity 78.6%; Pred. No. 3.1e+02; Mismatches 6; Indels 0; Gaps 0;
Matches 22; Conservative 0; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
Db 34 GTGAAGCAGTGACAGCATGTCGTCAGGC 61

; GENERAL INFORMATION:
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 492
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-492

Query Match          59.4%; Score 18.4; DB 15; Length 1799;
Best Local Similarity 78.6%; Pred. No. 3.1e+02; Mismatches 6; Indels 0; Gaps 0;
Matches 22; Conservative 0; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
Db 426 GTGAAGCAGTGACAGCATGTCGTCAGGC 399

RESULT 32
US-10-101-510-492/c
; Sequence 492, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 492
; LENGTH: 1799
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-492

Query Match          59.4%; Score 18.4; DB 15; Length 1799;
Best Local Similarity 78.6%; Pred. No. 3.1e+02; Mismatches 6; Indels 0; Gaps 0;
Matches 22; Conservative 0; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
Db 426 GTGAAGCAGTGACAGCATGTCGTCAGGC 399

RESULT 33
US-10-412-699B-403/c
; Sequence 403, Application US/10412699B
; Publication No. US20040045049A1
; GENERAL INFORMATION:
; APPLICANT: Mendel Biotechnology, Inc.
; APPLICANT: Zhang, James
; APPLICANT: Fromm, Michael E.
; APPLICANT: Heard, Jacqueline E.
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Adam, Luc J.
; APPLICANT: Broun, Pierre E.
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James S.
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Samaha, Raymond R.
; APPLICANT: Pilgrim, Marsha L.
; APPLICANT: Creelman, Robert A.
; APPLICANT: DuBell, Arnold N.
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Kumimoto, Roderick
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: Polynucleotides and Polypeptides in Plants
; FILE REFERENCE: MBI-0048CIP
; CURRENT APPLICATION NUMBER: US/10/412,699B
; CURRENT FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: 09/394,519
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: 09/489,376
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/506,720
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 09/533,030
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,392
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,029
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/532,591
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/533,648
```

; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 09/713,994
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 09/819,142
; PRIOR FILING DATE: 2001-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 403
; LENGTH: 2291
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G624
US-10-412-699B-403

Query Match 59.4%; Score 18.4; DB 17; Length 2291;
Best Local Similarity 78.6%; Pred. No. 3.2e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 TGAAGCGGTGACGACGATATCAGCGCT 31
DB 1424 TCAAGCGGTGAATGCATATCAGCGGAT 1397

RESULT 34
US-10-108-260A-1214
; Sequence 1214, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1214
; LENGTH: 3056
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1214

Query Match 59.4%; Score 18.4; DB 17; Length 3056;
Best Local Similarity 78.6%; Pred. No. 3.3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 3 GTGAAGCGGTGACGACGATATCAGCGGC 30
DB 1499 GAGAAGGGGTGGACCATATGAACGGC 1526

RESULT 35
US-10-128-714-223/c
; Sequence 223, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wengli
; APPLICANT: Tienkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and Methods of Use
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890

; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
; LENGTH: 3826
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-223

Query Match 59.4%; Score 18.4; DB 15; Length 3826;
Best Local Similarity 78.6%; Pred. No. 3.3e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 TGAAGCGGTGACGACGATATCAGCGCT 31
DB 3447 TGAAGCGGTGACGACGCTCTCCGACTTCT 3420

RESULT 36
US-10-437-963-10983/c
; Sequence 10983, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 10983
; LENGTH: 4473
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_17250C.1
US-10-437-963-10983

Query Match 59.4%; Score 18.4; DB 18; Length 4473;
Best Local Similarity 95.0%; Pred. No. 3.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 9 CGGTGACGACGATATCAGACG 28
DB 4163 CGGTGACGACGATATCAGACG 4144

RESULT 37
US-10-278-536-223/c
; Sequence 223, Application US/10278536
; Publication No. US20030131386A1
; GENERAL INFORMATION:
; APPLICANT: Samaha, Raymond
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Pineda, Omaira
; APPLICANT: Reuber, Lynne
; APPLICANT: Riechmann, Jose-Luis
; APPLICANT: Yu, Guo-Liang
; APPLICANT: Keddie, James
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Pilgrim, Marsha

APPLICANT: Adam, Luc
APPLICANT: Broun, Pierre
TITLE OF INVENTION: STRESS-INDUCED POLYNUCLEOTIDES
FILE REFERENCE: MBI-011
CURRENT APPLICATION NUMBER: US/10/278,536
PRIOR FILING DATE: 2002-10-22
PRIOR APPLICATION NUMBER: 60/125,814
PRIOR FILING DATE: 1999-03-23
NUMBER OF SEQ ID NOS: 238
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 223
LENGTH: 5394
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G624
US-10-278-536-223

Query Match 59.4%; Score 18.4; DB 15; Length 5394;
Best Local Similarity 78.6%; Pred. No. 3.5e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 TGAAGCGGTGACAGCATATCAGCGCT 31
Db 3987 TCAACGGTGAATGCATATCAGCCGAT 3960

RESULT 38
US-09-984-429-394
Sequence 394, Application US/09984429
Publication No. US20040010132A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 53 Human Secreted Proteins
FILE REFERENCE: PZ018P2
CURRENT APPLICATION NUMBER: US/09/984,429
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,591
PRIOR FILING DATE: 2000-11-01
PRIOR FILING DATE: 1999-04-08
PRIOR APPLICATION NUMBER: 60/061,527
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,463
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,529
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/071,498
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,532
PRIOR FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 727
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 394
LENGTH: 11122
TYPE: DNA
ORGANISM: Homo sapiens
US-09-984-429-394

Query Match 59.4%; Score 18.4; DB 11; Length 11122;
Best Local Similarity 78.6%; Pred. No. 3.7e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGCGC 30
Db 6416 GTGAAGGAAGCCAGCATATCAGATGGC 6443

RESULT 39

US-09-984-429-444
Sequence 444, Application US/09984429
Publication No. US20040010132A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 53 Human Secreted Proteins
FILE REFERENCE: PZ018P2
CURRENT APPLICATION NUMBER: US/09/984,429
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: 60/244,591
PRIOR FILING DATE: 2000-11-01
PRIOR FILING DATE: 1999-04-08
PRIOR APPLICATION NUMBER: 60/061,463
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,529
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/071,498
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,527
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,536
PRIOR FILING DATE: 1997-10-09
PRIOR APPLICATION NUMBER: 60/061,532
PRIOR FILING DATE: 1997-10-09
NUMBER OF SEQ ID NOS: 727
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 444
LENGTH: 11122
TYPE: DNA
ORGANISM: Homo sapiens
US-09-984-429-444

Query Match 59.4%; Score 18.4; DB 11; Length 11122;
Best Local Similarity 78.6%; Pred. No. 3.7e+02;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGCGC 30
Db 6416 GTGAAGGAAGCCAGCATATCAGATGGC 6443

RESULT 40
US-09-070-927A-49
Sequence 49, Application US/09070927A
Patent No. US20020120116A1
GENERAL INFORMATION:
APPLICANT: Charles A. Kunsch
Steven Barash
Patrick J. Dillon
TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 982
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,927A
FILING DATE: 04-May-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/046,655
FILING DATE: 1997-05-16

```

; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11307 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-070-927A-49

```

```

Query Match      59.4%; Score 18.4; DB 9; Length 11307;
Best Local Similarity 73.3%; Pred. No. 3.7e+02;
Matches 22; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      1 TACTGAGCGGTGACAGCATATCAGACGGC 30
Db      9635 TGGTGAAGYCGAGGCAAAATATCAGCGGC 9664

```

Search completed: June 4, 2005, 12:20:16
Job time : 304.509 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:32:00 ; Search time 1892.65 Seconds
(without alignments)
623.460 Million cell updates/sec

Title: US-09-674-277-26

Perfect score: 31
Sequence: 1 tagtgaagcgtgacagcatatcagacggc 31

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gse1: *
9: gb_gse2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	67.7	660	9	AL095079 Arabidops
2	20.2	65.2	1135	9	AL255773 Tetraodon
C 3	20	64.5	261	8	AZ536456 110300.50
C 4	20	64.5	436	1	A1544961 fb69h06.y
5	20	64.5	467	1	AA158547 zo72g09.x
6	20	64.5	669	6	CA588763 hab68d06.
C 7	20	64.5	789	6	CK142335 AGENCOURT
8	20	64.5	890	6	CD558197 AGENCOURT
C 9	20	64.5	1157	8	B10296 F13J13-Sp6
C 10	19.8	63.9	374	7	CK087334 A003P70.3
C 11	19.8	63.9	424	7	CK114607 X043A02.P
C 12	19.8	63.9	559	8	AQ759220 HS_2206.B
C 13	19.8	63.9	883	6	CA791117 AGENCOURT
C 14	19.8	63.9	889	9	CG164810 PUJEU04TD
C 15	19.8	63.9	897	8	BZ747628 PUJEU51TD
C 16	19.8	63.9	911	8	CC011362 PUDPY96TD
C 17	19.8	63.9	1319	5	BU414461 603668825
C 18	19.6	63.2	343	5	BP041566 BP041566
C 19	19.6	63.2	508	4	BG385140 306760.MA
C 20	19.6	63.2	801	9	CG035531 PUH497B
C 21	19.6	63.2	898	9	CL240413 ZMMBB05B
C 22	19.4	62.6	279	8	AQ193913 CIT-HSP-2
C 23	19.4	62.6	306	6	CB239270 TGESTzyf5
C 24	19.4	62.6	348	7	CO743178 TGESTzyf3

C 25	19.4	62.6	451	7	CN122238
C 26	19.4	62.6	462	2	AW895689
C 27	19.4	62.6	479	7	CN195909
C 28	19.4	62.6	506	7	CN620163
C 29	19.4	62.6	534	7	CF341248
C 30	19.4	62.6	541	7	CF247429
C 31	19.4	62.6	545	6	CD217441
C 32	19.4	62.6	546	4	BM176352
C 33	19.4	62.6	549	4	BM189407
C 34	19.4	62.6	556	4	BM189147
C 35	19.4	62.6	562	7	CB268779
C 36	19.4	62.6	567	6	CB301963
C 37	19.4	62.6	574	6	CB187210
C 38	19.4	62.6	576	7	CN619263
C 39	19.4	62.6	578	7	CB187908
C 40	19.4	62.6	585	6	CB382287
C 41	19.4	62.6	587	7	CN617452
C 42	19.4	62.6	592	7	CN614478
C 43	19.4	62.6	594	4	BM176326
C 44	19.4	62.6	599	1	AL887114
C 45	19.4	62.6	603	7	CN621196

ALIGNMENTS

RESULT 1
CNS000MXM
LOCUS
DEFINITION
CNS000MXM 660 bp DNA linear GSS 29-JUN-1999
Arabidopsis thaliana genome survey sequence T7 end of BAC T14P16 of
TAMU library from strain Columbia of Arabidopsis thaliana, genomic
survey sequence.
ACCESSION
AL095079
VERSION
AL095079.1 GI:5303234
KEYWORDS
GSS.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
REFERENCE
1 (bases 1 to 660)
AUTHORS
Salamounat, M., Choise, N., Artiguenave, P., Brottier, P., Wincker, P.,
Samson, D., Saurin, W., Weissenbach, J. and Quetier, P.
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 660)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

FEATURES

source
1..660
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="T14P16"
/note="end : T7"

ORIGIN

Query Match 67.7%; Score 21; DB 9; Length 660;
Best Local Similarity 82.8%; Pred. No. 1.7e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 ACTGAGCGGTGACAGCATATCAGACGGC 30
DB 303 ACTGTAGCGGTGCCAGATAGCAGCGGC 331

RESULT 2

CNS03QCK
LOCUS
CNS03QCK 1135 bp DNA linear GSS 01-SEP-2000

DEFINITION
Tetraodon nigroviridis genome survey sequence T7 end of clone 047G05 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION
AL255773

VERSION
AL255773.1 GI:7976785

KEYWORDS
GSS; genome survey sequence.

SOURCE
Tetraodon nigroviridis

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE
1 Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, P., Saurin, W. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633
10835645

REFERENCE
2

AUTHORS
Roest Crolius, H., Jaillon, O., Dasilva, C., Orzouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, P., Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
20359837
10899143

MEDLINE
PUBMED

REFERENCE
3 (bases 1 to 1135)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr]

TITLE
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
source
1. .1135
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="047G05"
/clone_lib="G"
/note="Genoscope sequence ID : COBG047AD03LP1-end : T7"

ORIGIN
Query Match 65.2%; Score 20.2; DB 9; Length 1135;
Best Local Similarity 88.0%; Pred. No. 4.1e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY
2 ACTGAAGCGGTGACGATATCAGA 26
|||||

Db
619 ACTGAAGCGGTACAGCATGACAGA 643
|||||

RESULT 3
AZ536456/c

LOCUS
110300.50 Planococcus lilacinus DNA linear GSS 03-NOV-2000

DEFINITION
genomic survey sequence.

ACCESSION
AZ536456

VERSION
AZ536456.1 GI:11093403

KEYWORDS
GSS.

SOURCE
Planococcus lilacinus (lilac mealybug)

ORGANISM
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes; Coccoidea; Pseudococcidae; Planococcus.

REFERENCE
1 (bases 1 to 261)

AUTHORS
Mohan, K.N. and Chandra, H.S.
Mealybug shotgun sequencing
Unpublished (2000)
Contact: Mohan KN
Microbiology and Cell Biology
Indian Institute of Science
Sir C.V. Raman Avenue, Bangalore, Karnataka 560012, India
Email: mohan@cbl.iisc.ernet.in
Class: shotgun.

FEATURES
source
1. .261
/organism="Planococcus lilacinus"
/mol_type="genomic DNA"
/db_xref="taxon:40930"
/clone_lib="Planococcus lilacinus DNA"

ORIGIN
Query Match 64.5%; Score 20; DB 8; Length 261;
Best Local Similarity 82.1%; Pred. No. 4.2e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY
4 TCAAGCGGTGACGATATCAGCGCT 31
|||||

Db
243 TTAAGGGTAGACATCATATCAGCGCT 216
|||||

RESULT 4
AI544961/c

LOCUS
AI544961.1

DEFINITION
EST.

ACCESSION
AI544961

VERSION
AI544961.1 GI:4462334

KEYWORDS
EST.

SOURCE
Danio rerio (zebrafish)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 436)
Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by: Matthew Clark. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Genome Systems, St. Louis, Missouri (web address: www.genomesystems.com) (email contact: info@genomesystems.com) and Research Genetics, Huntsville, Alabama (web address: www.resgen.com) (email contact: info@resgen.com) and Resourcenzentrum Primatdatenbank, Berlin, Germany (web address: www.rzp.de)
Possible reversed clone: similarity on wrong strand
zebrafish identity [p-value greater than 1e-99] found to:
gi|2289453|gb|AA542518|AA542518 fa07e01.r1 Zebrafish ICRPzfls Danio rerio cDNA
Seq primer: T3 ET from Amersham
High quality sequence stop: 324
POLYA=No.
Location/Qualifiers
1. .436
/organism="Danio rerio"

/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:3717179"
/sex="mixed"
/tissue_type="26 somite embryos, adult livers, shield stage embryos"
/lab_host="X11-blue MRP"
/clone_lib="zebrafish WashU MPMG EST"
/notes="vector: pSPORT1; Site 1: NotI; Site 2: SalI; 1st strand cDNA was primed with a Not I - oligo(drf)15 primer [5'pGACTAGTTCTAGATCGGAGCGCGCCCTTTTTTTTTTTT3']; double-stranded cDNA was ligated to Sal I adaptors (BRL), digested with Not I and cloned into the Not I and Sal I sites of the pSPORT1 vector (BRL). Library was constructed by Matthew Clark (Lehrach lab; ICRF, London and Max Planck Institut fuer Molekulare Genetik, Berlin). cDNAs for EST analysis were selected following oligonucleotide hybridization fingerprinting of arrayed clones from zebrafish late somitogenesis (26 ss), adult liver or embryonic shield stage (5.6 h) libraries. Fingerprint data were used to computationally cluster cDNAs, and a single cDNA from each cluster was chosen for sequencing. In some cases multiple members of the same cluster were sequenced to assess clustering parameters or single clones were sequenced additional times to assess quality control."

ORIGIN

Query Match 64.5%; Score 20; DB 1; Length 436;
Best Local Similarity 82.1%; Pred. No. 4.5e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGG 29
|||||
DB 379 AGTGAAGCGGTGCGAGCTGATCAGACCG 352
|||||

RESULT 5

AA158547 467 bp mRNA linear EST 16-DEC-1996
LOCUS z072q09.r1 StrataGene pancreas (#937208) Homo sapiens cDNA clone
DEFINITION IMAGE:592480 5' similar to TR:G556769 G556769 INOSITOL
1,4,5-TRIPHOSPHATE 5-PHOSPHATASE.1, mRNA sequence.

ACCESSION AA158547 GI:1733358
VERSION EST.
KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 467)

REFERENCE

AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiappelli,B., Chisoso,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,W.B., Tan,F., Thierry-Mieg,J., Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

TITLE

JOURNAL Washington University School of Medicine
MEDLINE 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
PUBMED Tel: 314 286 1800
8889549 Fax: 314 286 1810
Email: estowatson.wustl.edu

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 351.

FEATURES

Location/Qualifiers

source

1. .467
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4623866"
/db_xref="taxon:9606"
/clone="IMAGE:592480"
/lab_host="SOLAR cells (kanamycin resistant)"
/clone_lib="Stratagene pancreas (#937208)"
/note="Organ: pancreas; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'."

ORIGIN

Query Match 64.5%; Score 20; DB 1; Length 467;
Best Local Similarity 76.7%; Pred. No. 4.5e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGGCT 31
|||||
DB 95 AGTGATCGCATGAGAGATATNACAGGGCT 124
|||||

RESULT 6

CA588763 669 bp mRNA linear EST 19-NOV-2002
LOCUS hab8d06.y1 Fugu UT7 adult skin Takifugu rubripes cDNA clone
DEFINITION IMAGE:6353699 5' similar to TR:O89994 O89994 HYPOTHETICAL 38.2 KD PROTEIN.; contains MER22.t3 MER22 repetitive element.; mRNA sequence.

ACCESSION CA588763 GI:25133341
VERSION EST.
KEYWORDS

SOURCE Takifugu rubripes (Fugu rubripes)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Takifugu.
1 (bases 1 to 669)

REFERENCE

AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,P., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished (1998)

TITLE

JOURNAL Washington University School of Medicine
COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu

Library materials provided by G. Elgar (UK MRC HGMP-RC) Library constructed by Drs. K. Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Watabe (University of Tokyo, Institute of Medical Science and Laboratory of Aquatic Molecular Biology and Biotechnology) DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Fugu clone distribution information can be found through the I.M.A.G.E. Consortium/LML, send email to: info@image.llnl.gov

Possible reversed clone: similarity on wrong strand

Seq primer: T3 ET from Amersham

High quality sequence stop: 474.

FEATURES

source

1. .669
/organism="Takifugu rubripes"
/mol_type="mRNA"
/db_xref="taxon:31033"
/clone="IMAGE:6353699"

```

/sex="female"
/tissue_type="skin"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="Fugu UT7 adult skin"
/notes="Vector: pBluescript-FL; Site 1: pFLM I
(CCANNNNTGG); Site 2: pFLM I (CCANNNNTGG); BamHI-SmaI
sites were converted to BamHI-PfIMI-SfiI-PfIMI sites (SmaI
is destroyed). Other part of the vector is untouched. The
cDNA is inserted between two pFLM I sites in T3 (5') to T7
(3') direction. Library materials provided by G. Elgar (UK
MRC HGMP-RC) and constructed and donated by Drs. K.
Kawakami, M. Sasaki, S. Sugano, K. Kikuchi and S. Watabe
(University of Tokyo, Institute of Medical Science and
Laboratory of Aquatic Molecular Biology and
Biotechnology)."

```

ORIGIN

```

Query Match      64.5%; Score 20; DB 6; Length 669;
Best Local Similarity 82.1%; Pred. No. 4.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TGAAGCGGTGACAGCATATCAGACGGCT 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 224 TGAAGCGGTGACAGCATATTTGCCGGCT 251

```

RESULT 7

```

CK142335/c
LOCUS CK142335
DEFINITION AGENCOURT_16821868 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7048896
5', mRNA sequence.
ACCESSION CK142335
VERSION CK142335.1 GI:38647532
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

```

REFERENCE

```

AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14817 row: e column: 22
High quality sequence stop: 379.
Location/Qualifiers
1..789
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7048896"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available

```

FEATURES

source

```

Query Match      64.5%; Score 20; DB 6; Length 669;
Best Local Similarity 82.1%; Pred. No. 4.7e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 TGAAGCGGTGACAGCATATCAGACGGCT 31
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 224 TGAAGCGGTGACAGCATATTTGCCGGCT 251

```

```

(NIH ZGC 7). Library was constructed by Open Biosystems
(Huntsville, AL)."

```

ORIGIN

```

Query Match      64.5%; Score 20; DB 7; Length 789;
Best Local Similarity 82.1%; Pred. No. 4.8e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AGTGAAGCGGTGACAGCATATCAGACGG 29
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 553 AGTGAAGCGGTGACAGCATATCAGACGG 526

```

RESULT 8

```

CD558197
LOCUS CD558197
DEFINITION AGENCOURT_14477493 NIH_MGC_181 Homo sapiens cDNA clone
IMAGE:30395451 5', mRNA sequence.
ACCESSION CD558197
VERSION CD558197.1 GI:31584265
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

```

AUTHORS NIH-MGC http://mgs.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM475 row: j column: 04
High quality sequence stop: 635.
Location/Qualifiers
1..890
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30395451"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A ( T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_181"
/notes="Vector: pCMV-Sp6T6.1; Site 1: NotI; Site 2: EcoRV
(deleted); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."

```

FEATURES

source

```

Query Match      64.5%; Score 20; DB 6; Length 890;
Best Local Similarity 82.1%; Pred. No. 4.9e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 418 GTGAAGCGGTGACAGCATATTCAGAAGGC 445

```

ORIGIN

```

Query Match      64.5%; Score 20; DB 6; Length 890;
Best Local Similarity 82.1%; Pred. No. 4.9e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGACGGC 30
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 418 GTGAAGCGGTGACAGCATATTCAGAAGGC 445

```

RESULT 9

```

B10296/c
LOCUS B10296
DEFINITION F13J13-Sp6 IGF Arabidopsis thaliana genomic clone F13J13, genomic
1157 bp DNA linear GSS 14-MAY-1997

```

```

survey sequence.
ACCESSION      B10296
VERSION        B10296.1  GI:2091416
KEYWORDS       GSS.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
                Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE      1 (bases 1 to 1157)
AUTHORS        Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and
                Ecker, J.
TITLE          BAC End Sequences at ATGC
JOURNAL         Unpublished (1997)
COMMENT        Other_GSSs: F13J13-T7
                Contact: Ecker J.
                Arabidopsis Thaliana Genome Center
                University of Pennsylvania
                Dept. of Biology, University of Pennsylvania, Philadelphia, PA
                19104
                Tel: 215-898-9384
                Fax: 215-898-8780
                Email: jecker@atgenome.bio.upenn.edu
                Seq primer: Sp6
                Class: BAC ends
                High quality sequence start: 66
                High quality sequence stop: 671.
                Location/Qualifiers
                1..1157
                /organism="Arabidopsis thaliana"
                /mol_type="genomic DNA"
                /scot_type="Columbia"
                /db_xref="taxon:3702"
                /clone="F13J13"
                /sex="hermaphrodite"
                /clone_lib="IGF"
                /notes="Vector: BelOBACII; Site_1: EcoRI; Site_2: EcoRI;
                Produced by Thomas Altman"

ORIGIN
Query Match      64.5%; Score 20; DB 8; Length 1157;
Best Local Similarity 82.1%; Pred. No. 5.1e+02;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      3 GTGAAGCGGTGACGACATATCAGCGGC 30
        ||||||| ||||||| |||||||
DB      817 GAGAAGCGGAGACAGCATAGAGCGGC 790

RESULT 10
CK087334/c
LOCUS          CK087334 374 bp mRNA linear EST 01-DEC-2003
DEFINITION    A003P70.3pR Hybrid aspen plasmid library Populus tremula x Populus
                tremuloides cDNA clone A003P70 3', mRNA sequence.
ACCESSION     CK087334
VERSION       CK087334.1 GI:38571548
KEYWORDS      EST.
SOURCE        Populus tremula x Populus tremuloides
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 374)
AUTHORS        Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,
                Brunner, A.M., Campaa, L., Jonsson-Lindvall, J., Tandré, K.,
                Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P.,
                Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J. and Jansson, S.
                A Populus EST resource for functional genomics
                Unpublished (2003)
                Contact: Bo Segerman
                Umea Plant Science Center, Department of Plant Physiology
                Umea University
                901 87 Umea, Sweden
                Tel: +46 90 786 5279
                Fax: +46 90 786 6676
                Email: bo.segerman@plantphys.umu.se.
                Location/Qualifiers
                1..424
                /organism="Populus tremula x Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:47664"
                /clone="X043A02"
                /tissue_type="wood"
                /clone_lib="Populus wood cDNA library"

TITLE          A Populus EST resource for functional genomics
JOURNAL         Unpublished (2003)
COMMENT        Other ESTs: A003P70U
                Contact: Bo Segerman
                Umea Plant Science Center, Department of Plant Physiology
                Umea University
                901 87 Umea, Sweden
                Tel: +46 90 786 5279
                Fax: +46 90 786 6676
                Email: bo.segerman@plantphys.umu.se.
                Location/Qualifiers
                1..424
                /organism="Populus tremula x Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:47664"
                /clone="X043A02"
                /tissue_type="wood"
                /clone_lib="Populus wood cDNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 7; Length 424;
Best Local Similarity 77.4%; Pred. No. 5.5e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 TAGTGAAGCGGTGACGACATATCAGCGGCT 31
        ||||||| ||||||| |||||||
DB      56 TAGTGAAGAGCTGGCAGTATACAGACTGTT 26

RESULT 11
CK114607
LOCUS          CK114607 424 bp mRNA linear EST 01-DEC-2003
DEFINITION    X043A02 Populus wood cDNA library Populus tremula x Populus
                tremuloides cDNA clone X043A02 5', mRNA sequence.
ACCESSION     CK114607
VERSION       CK114607.1 GI:38598932
KEYWORDS      EST.
SOURCE        Populus tremula x Populus tremuloides
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 424)
AUTHORS        Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,
                Brunner, A.M., Campaa, L., Jonsson-Lindvall, J., Tandré, K.,
                Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P.,
                Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J. and Jansson, S.
                A Populus EST resource for functional genomics
                Unpublished (2003)
                Contact: Bo Segerman
                Umea Plant Science Center, Department of Plant Physiology
                Umea University
                901 87 Umea, Sweden
                Tel: +46 90 786 5279
                Fax: +46 90 786 6676
                Email: bo.segerman@plantphys.umu.se.
                Location/Qualifiers
                1..424
                /organism="Populus tremula x Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:47664"
                /clone="X043A02"
                /tissue_type="wood"
                /clone_lib="Populus wood cDNA library"

TITLE          A Populus EST resource for functional genomics
JOURNAL         Unpublished (2003)
COMMENT        Other ESTs: A003P70U
                Contact: Bo Segerman
                Umea Plant Science Center, Department of Plant Physiology
                Umea University
                901 87 Umea, Sweden
                Tel: +46 90 786 5279
                Fax: +46 90 786 6676
                Email: bo.segerman@plantphys.umu.se.
                Location/Qualifiers
                1..424
                /organism="Populus tremula x Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:47664"
                /clone="X043A02"
                /tissue_type="wood"
                /clone_lib="Populus wood cDNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 7; Length 424;
Best Local Similarity 77.4%; Pred. No. 5.5e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

901 87 Umea, Sweden
Tel: +46 90 786 5279
Fax: +46 90 786 6676
Email: bo.segerman@plantphys.umu.se.
Location/Qualifiers
1..374
/organism="Populus tremula x Populus tremuloides"
/mol_type="mRNA"
/db_xref="taxon:47664"
/clone="A003P70"
/tissue_type="Cambial region"
/dev_stage="1.5 m actively growing tree"
/lab_host="E.coli"
/clone_lib="Hybrid aspen plasmid library"
/notes="Vector: pBluescript SK; Site_1: SalI; Site_2: NotI;
Cambial region tissues, including developing xylem, the
meristematic cambial zone and the developing and mature
phloem, was harvested and cloned into lambda gt22a. DNA was
cDNA was prepared and cloned into lambda gt22a. DNA was
isolated and subcloned into pBluescript SK using SalI and
NotI restriction enzymes."

ORIGIN
Query Match      63.9%; Score 19.8; DB 7; Length 374;
Best Local Similarity 77.4%; Pred. No. 5.4e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      1 TAGTGAAGCGGTGACGACATATCAGCGGCT 31
        ||||||| ||||||| |||||||
DB      56 TAGTGAAGAGCTGGCAGTATACAGACTGTT 26

RESULT 11
CK114607
LOCUS          CK114607 424 bp mRNA linear EST 01-DEC-2003
DEFINITION    X043A02 Populus wood cDNA library Populus tremula x Populus
                tremuloides cDNA clone X043A02 5', mRNA sequence.
ACCESSION     CK114607
VERSION       CK114607.1 GI:38598932
KEYWORDS      EST.
SOURCE        Populus tremula x Populus tremuloides
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                rosids; eurosids I; Malpighiales; Salicaceae; Populus.
REFERENCE      1 (bases 1 to 424)
AUTHORS        Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P.,
                Brunner, A.M., Campaa, L., Jonsson-Lindvall, J., Tandré, K.,
                Strauss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P.,
                Nilsson, O., Sandberg, G., Karlsson, J., Lundberg, J. and Jansson, S.
                A Populus EST resource for functional genomics
                Unpublished (2003)
                Contact: Bo Segerman
                Umea Plant Science Center, Department of Plant Physiology
                Umea University
                901 87 Umea, Sweden
                Tel: +46 90 786 5279
                Fax: +46 90 786 6676
                Email: bo.segerman@plantphys.umu.se.
                Location/Qualifiers
                1..424
                /organism="Populus tremula x Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:47664"
                /clone="X043A02"
                /tissue_type="wood"
                /clone_lib="Populus wood cDNA library"

TITLE          A Populus EST resource for functional genomics
JOURNAL         Unpublished (2003)
COMMENT        Other ESTs: A003P70U
                Contact: Bo Segerman
                Umea Plant Science Center, Department of Plant Physiology
                Umea University
                901 87 Umea, Sweden
                Tel: +46 90 786 5279
                Fax: +46 90 786 6676
                Email: bo.segerman@plantphys.umu.se.
                Location/Qualifiers
                1..424
                /organism="Populus tremula x Populus tremuloides"
                /mol_type="mRNA"
                /db_xref="taxon:47664"
                /clone="X043A02"
                /tissue_type="wood"
                /clone_lib="Populus wood cDNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 7; Length 424;
Best Local Similarity 77.4%; Pred. No. 5.5e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
|||||
Db 359 TAGTGAAGAGCTGGCAGTATTAACAGACTGTT 389

RESULT 12
AQ759220/c
LOCUS HS_2206_B2_A02_MR_CIT_Approved_Human_Genomic_Sperm_Library_D_Homo
DEFINITION sapiens genomic clone Plate=2206 Col=4 Row=B, genomic survey
sequence.
ACCESSION AQ759220
VERSION AQ759220.1 GI:5624503
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 559)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end web Server: http://www.htsc.washington.edu
Plate: 2206 row: B column: 4
Seq primer: M13 Reverse
Class: BAC ends
High quality sequence stop: 559.
FEATURES
source
Location/Qualifiers
1..559
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=2206 Col=4 Row=B"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBeoBAC11; BAC Clones in
E-Coli DH10B"

ORIGIN
Query Match 63.9%; Score 19.8; DB 8; Length 559;
Best Local Similarity 77.4%; Pred. No. 5.7e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
|||||
Db 542 TAGTGATGCAGTAATAGTATATTAGACGCTCT 512

RESULT 13
CA791117
LOCUS AGENCOURT_10305625_NICHG_XGC_O01_Xenopus_laevis_cdna_clone
DEFINITION IMAGE:5075608 5', mRNA sequence.
ACCESSION CA791117
VERSION CA791117.1 GI:26037680
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 883)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11213 row: h column: 09
High quality sequence stop: 707.
FEATURES
source
Location/Qualifiers
1..883
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:5079608"
/tissue_type="ocytes"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHG_XGC_O01"
/note="Vector: pCHV-SFOR6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.2 kb. Constructed by Life Technologies."

ORIGIN
Query Match 63.9%; Score 19.8; DB 6; Length 883;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
|||||
Db 560 TACTGAAGCGTTTACAGTATATCTGAGGCTCT 590

RESULT 14
CG164810
LOCUS PUJEU04TD_ZM_0.6_1.0_KB_Zea_mays_genomic_clone_ZMMBTa0652A07,
DEFINITION genomic survey sequence.
ACCESSION CG164810
VERSION CG164810.1 GI:34055611
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 889)
AUTHORS Whetlaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Other_GSSs: PUJEU04TB
Contact: Cathy Whetlaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whetlaw@tigr.org
Seq primer: Tg
Class: sheared ends.
FEATURES
source
Location/Qualifiers
1..889
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"

```

```

/db_xref="taxon:4577"
/clone="ZMBF0652A07"
/clone_lib="ZM 0.6 1.0 KB"
/notes="Vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
Cot selected genomic DNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 9; Length 889;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
    ||||| ||||| ||||| ||||| ||||| |||||
Db 580 TAATGAATGCTGGCAGCAATCAGACGGCT 610

RESULT 15
BZ747628/c
LOCUS      BZ747628      897 bp      DNA      linear      GSS 03-MAR-2003
DEFINITION PUCQE51TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBFta129105,
            genomic survey sequence.
ACCESSION  BZ747628
VERSION    BZ747628.1 GI:28727338
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (Bases 1 to 897)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Contact: Cathy Whitelaw
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
            Location/Qualifiers
                source
                    1..911
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone_lib="ZM 0.6 1.0 KB"
                    /notes="Vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
                    Cot selected genomic DNA library"

ORIGIN
Query Match      53.9%; Score 19.8; DB 8; Length 911;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
    ||||| ||||| ||||| ||||| ||||| |||||
Db 54 TAGGAGGCAGTGTAGGACATCTAGGACGGCT 24

RESULT 17
BZ414461
LOCUS      BZ414461      1319 bp      mRNA      linear      EST 29-NOV-2003
DEFINITION 603668825F1 CSEQRBL06 Gallus gallus cDNA clone CHEST609b7 5', mRNA
            sequence.
ACCESSION  BZ414461
VERSION    BZ414461.1 GI:25907132
KEYWORDS   EST.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            1 (Bases 1 to 1319)
REFERENCE  Boardman,P.B., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
            Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
            A Comprehensive Collection of Chicken cDNAs
            Curr. Biol. 12 (22), 1965-1969 (2002)
            22335534
            PUBMED 12445392
COMMENT    Contact: Simon Hubbard
            Department of Biomolecular Sciences
            University of Manchester Institute of Science and Technology
            (UMIST)
            PO Box 88, Manchester, M60 1QD, UK
            Tel: 01612008930
            Fax: 01612360409
            Email: Simon.Hubbard@umist.ac.uk.
            Location/Qualifiers
                source
                    1..1319
                    /organism="Gallus gallus"
                    /mol_type="mRNA"
                    /strain="Layer and broiler"
                    /db_xref="taxon:9031"

ORIGIN
Query Match      63.9%; Score 19.8; DB 8; Length 897;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
    ||||| ||||| ||||| ||||| ||||| |||||
Db 836 TAGGAGGCAGTGTAGGACATCTAGGACGGCT 806

RESULT 16
CC011362/c
LOCUS      CC011362      911 bp      DNA      linear      GSS 31-MAR-2003
DEFINITION PUDFY96TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBFta178p23,
            genomic survey sequence.
ACCESSION  CC011362
VERSION    CC011362.1 GI:29392538
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (Bases 1 to 897)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Contact: Cathy Whitelaw
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
            Location/Qualifiers
                source
                    1..897
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone_lib="ZM 0.6 1.0 KB"
                    /notes="Vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
                    Cot selected genomic DNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 8; Length 897;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
    ||||| ||||| ||||| ||||| ||||| |||||
Db 836 TAGGAGGCAGTGTAGGACATCTAGGACGGCT 806

RESULT 16
CC011362/c
LOCUS      CC011362      911 bp      DNA      linear      GSS 31-MAR-2003
DEFINITION PUDFY96TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBFta178p23,
            genomic survey sequence.
ACCESSION  CC011362
VERSION    CC011362.1 GI:29392538
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (Bases 1 to 897)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Contact: Cathy Whitelaw
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
            Location/Qualifiers
                source
                    1..897
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone_lib="ZM 0.6 1.0 KB"
                    /notes="Vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
                    Cot selected genomic DNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 8; Length 897;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
    ||||| ||||| ||||| ||||| ||||| |||||
Db 836 TAGGAGGCAGTGTAGGACATCTAGGACGGCT 806

RESULT 16
CC011362/c
LOCUS      CC011362      911 bp      DNA      linear      GSS 31-MAR-2003
DEFINITION PUDFY96TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBFta178p23,
            genomic survey sequence.
ACCESSION  CC011362
VERSION    CC011362.1 GI:29392538
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (Bases 1 to 897)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Contact: Cathy Whitelaw
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
            Location/Qualifiers
                source
                    1..897
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone_lib="ZM 0.6 1.0 KB"
                    /notes="Vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
                    Cot selected genomic DNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 8; Length 897;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
    ||||| ||||| ||||| ||||| ||||| |||||
Db 836 TAGGAGGCAGTGTAGGACATCTAGGACGGCT 806

RESULT 16
CC011362/c
LOCUS      CC011362      911 bp      DNA      linear      GSS 31-MAR-2003
DEFINITION PUDFY96TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBFta178p23,
            genomic survey sequence.
ACCESSION  CC011362
VERSION    CC011362.1 GI:29392538
KEYWORDS   GSS.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE  1 (Bases 1 to 897)
AUTHORS   Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
            Reenick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
            Bennetzen,J.
            Maize Genomics Consortium
            Unpublished (2003)
            Contact: Cathy Whitelaw
            Email: whitelaw@tigr.org
            Seq primer: TP
            Class: sheared ends.
            Location/Qualifiers
                source
                    1..897
                    /organism="Zea mays"
                    /mol_type="genomic DNA"
                    /strain="B73"
                    /db_xref="taxon:4577"
                    /clone_lib="ZM 0.6 1.0 KB"
                    /notes="Vector: PCR4-TOFO; Site_1: EcoRI; 0.6-1.0 kb high
                    Cot selected genomic DNA library"

ORIGIN
Query Match      63.9%; Score 19.8; DB 8; Length 897;
Best Local Similarity 77.4%; Pred. No. 6e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
    ||||| ||||| ||||| ||||| ||||| |||||
Db 836 TAGGAGGCAGTGTAGGACATCTAGGACGGCT 806

RESULT 16
CC011362/c
LOCUS      CC011362      911 bp      DNA      linear      GSS 31-MAR-2003
DEFINITION PUDFY96TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBFta178p23,
            genomic survey sequence.
ACCESSION  CC011362
VERSION    CC011362.1 GI:29392538

```

/clone="CHEST609b7"
 /sex="Male and female"
 /issue_type="Abdominal fat pad"
 /dev_stage="adult"
 /lab_host="DH108"
 /clone_lib="CSEORBL06"
 /notes="Vector: pBluescript II KS(+); Site 1: EcoRI;
 Site 2: NotI; Modification of pBluescript II KS(+)
 [stratagene] vector to accommodate cDNA produced with the
 T-trimmed protocol (Construction of uni-directionally
 cloned cDNA libraries from messenger RNA for improved 3'
 end DNA sequencing by Glenn Fu, et al. U.S. Patent #
 6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
 Ligate in double stranded adaptor containing BsgI and
 BamHI sites [5'ggcgcgtgcagcccgatccgaaaaag
 [5'aattcttttttcgatacggggtcgacgc]

ORIGIN
 Query Match 63.9%; Score 19.8; DB 5; Length 1319;
 Best Local Similarity 77.4%; Pred. No. 6.3e+02;
 Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACAGCATATCAGACGGCT 31
 |||||
 Db 869 TTGTGACCGCATATAGTAGTACGCGCT 899

RESULT 18
 BP041566/c
 LOCUS
 DEFINITION
 BP041566 Lotus corniculatus var. japonicus flower bud Lotus
 corniculatus var. japonicus cDNA clone MFBL006b06_f 3', mRNA
 sequence.
 BP041566
 BP041566.1 GI:45573439
 EST.
 SOURCE
 ORGANISM
 Lotus corniculatus var. japonicus (Lotus japonicus)
 Lotus corniculatus var. japonicus
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
 Lotus.
 1 (bases 1 to 343)
 Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 Characteristics of the Lotus japonicus Gene Repertoire Deduced from
 Large-Scale Expressed Sequence Tag (EST) Analysis
 Plant Mol. Biol. 54 (3), 405-414 (2004)
 Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1..343
 /organism="Lotus corniculatus var. japonicus"
 /mol_type="mRNA"
 /isolate="Miyakojima MG-20"
 /db_xref="taxon:34305"
 /clone="MFBL006b06_f"
 /issue_type="flower bud"
 /clone_lib="Lotus corniculatus var. japonicus flower bud"

ORIGIN
 Query Match 63.2%; Score 19.6; DB 5; Length 343;
 Best Local Similarity 84.6%; Pred. No. 6.6e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 ACTGAAGCGGTGACAGCATATCAGAC 27
 |||||
 Db 324 ACTGAAGTGGTGAAGACATTTGAGAC 299

RESULT 19

BG385140
 LOCUS
 DEFINITION
 BG385140 306760 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BG385140.1 GI:13309612
 EST.
 Sus scrofa (pig)
 Sus scrofa
 Sus scrofa
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 508)
 Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
 Vallet, J., Wise, T., Rohrer, G.A., Perle, G., Sultana, R.,
 Quackenbush, J. and Keele, J.W.
 Porcine gene discovery by normalized cDNA-library sequencing and
 EST cluster assembly
 Mamm. Genome 13 (8), 475-478 (2002)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGCG
 Plate: 94 row: L column: 3
 Seq primer: ATTTAGTGACACTATAG.
 Location/Qualifiers
 1..508
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /issue_type="pooled"
 /lab_host="DH108"
 /clone_lib="MARC 1PIG"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

FEATURES
 source

ORIGIN

Query Match 63.2%; Score 19.6; DB 4; Length 508;
 Best Local Similarity 84.6%; Pred. No. 6.9e+02;
 Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 TGAAGCGGTGACAGCATATCAGACGG 29
 |||||
 Db 201 TGGAGCGGTGACAGCACTCAGAGG 226

RESULT 20
 CG035531
 LOCUS

DEFINITION
 CG035531 PUIHW49TB ZM_0.6_1.0_KB Zea mays genomic clone ZMMBTa0587I01,
 genomic survey sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CG035531.1 GI:33907687
 GSS.
 Zea mays
 Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 801)
 Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
 Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
 Benetzen, J.
 Maize Genomics Consortium

TITLE

JOURNAL Unpublished (2003)
COMMENT Other_GSSs: FUIHW49TD
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.

FEATURES
source Location/Qualifiers
1. .801
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBb0581J02"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: PCR1-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN
Query Match 63.2%; Score 19.6; DB 9; Length 801;
Best Local Similarity 84.6%; Pred. No. 7.3e+02;
Matches: 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACACATATCAGACG 28
Db 256 GGGAGCGGTGACACATGAGG 281

RESULT 21
CL240413/c
LOCUS
DEFINITION ZMMBb0581J02r ZMMBb (HindIII) Zea mays genomic clone
ACCESSION CL240413
VERSION CL240413.1 GI:40899251
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 898)
AUTHORS Bharti,A.K., Young,S., Kavchok,S., Keizer,G., Bronzino,A.C.,
Zohovetz,V., Fuks,G., Yu,Y., Wing,R. and Messing,J.
TITLE Sequencing of the maize genome at PGIR (2003c)
JOURNAL Unpublished (2003)
COMMENT Contact: Bharti,A.K.
Dr.Joachim Messing's lab
The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers
University
190 Frelinghuysen Road, Piscataway, NJ 08854, USA
Tel: 732 445 3801
Fax: 732 445 5735
Email: bharti@waksman.rutgers.edu
Seq primer: SP6
Class: BAC ends
High quality sequence start: 401.

FEATURES
source Location/Qualifiers
1. .898
/organism="Zea mays"
/mol_type="genomic DNA"
/cultivar="B73"
/db_xref="taxon:4577"
/clone="ZMMBb0581J02"
/lab_host="E. coli DH10B"
/clone_lib="ZMMBb (HindIII)"
/note="Vector: pCUGI; Site_1: HindIII; Site_2: HindIII"

ORIGIN
Query Match 63.2%; Score 19.6; DB 9; Length 898;
Best Local Similarity 84.6%; Pred. No. 7.3e+02;
Matches: 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Best Local Similarity 84.6%; Pred. No. 7.4e+02;
Matches: 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACACATATCAGACG 28
Db 549 GGGAGCGGTGACACATGAGG 524

RESULT 22
AQ193913
LOCUS
DEFINITION CIT-HSP-2384L11.TF CIT-HSP Homo sapiens genomic clone 2384L11,
genomic survey sequence.
ACCESSION AQ193913
VERSION AQ193913.1 GI:3605525
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 279)
AUTHORS Adams,M.D., Rounale,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source Location/Qualifiers
1. .279
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2384L11"
/sex="Male"
/cell_type="Sperm"
/clone_lib="CIT-HSP"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

ORIGIN
Query Match 62.6%; Score 19.4; DB 8; Length 279;
Best Local Similarity 79.3%; Pred. No. 7.9e+02;
Matches: 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TAGTGAAGCGGTGACACATATCAGACG 29
Db 191 TAGAGAAATTGGTGAACCATATGAGAGG 219

RESULT 23
CB239270/c
LOCUS
DEFINITION TgESTzf56h09.y1 TgrH Tachyzoite Norm 7 cDNA Library Toxoplasma
gondii cDNA clone TgESTzf56h09.y1 5', similar to SW:TB_P1AFK
P14643 TUBULIN BETA CHAIN. ;, mRNA sequence.
ACCESSION CB239270
VERSION CB239270.1 GI:28315814
KEYWORDS EST.
SOURCE Toxoplasma gondii
ORGANISM Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Email: toxo@watson.wustl.edu
 Contact David Sibley (toxost@borcim.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 407.

FEATURES

Location/Qualifiers
 1. .451
 /organism="Toxoplasma gondii"
 /mol_type="mRNA"
 /strain="VEG"
 /db_xref="taxon:5811"
 /clone="TgESTzyl22a04.y1"
 /dev_stage="Tachyzoite"
 /lab_host="GC10"
 /clone_lib="TgVEG118 Tachyzoite cDNA Library-2"

/note="Vector: pBluescript SK; Site 1: EcoRI; Site 2:
 XhoI; The library was constructed by Keliang Tang, Robert
 Cole and L. David Sibley at Washington University. cDNAs
 were synthesized from poly(A)+ RNAbly oligo d(T) priming,
 size-selected and directionally cloned into the Uni-ZAP XR
 lambda vector (Stratagene). The primary library was mass
 excised as phagemids and rescued in SOLR cells. The
 plasmid library was recovered from the SOLR cells and
 transformed in mass into GC10 cells for sequencing.
 WARNING: the library may contain a small percentage
 contaminants from human fibroblast cells."

ORIGIN

Query Match 62.6%; Score 19.4; DB 7; Length 451;
 Best Local Similarity 79.3%; Pred. No. 8.4e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAACGGGTGACGATATCAGCGCT 31
 ||||| ||||| ||||| ||||| |||||
 Db 303 GCGAAGCGGAGACGAGCGGTCAACCGCT 275

RESULT 26
 AW895689/c
 LOCUS
 DEFINITION QV4-NN0039-290300-154-a01 NN0039 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW895689
 VERSION AW895689.1 GI:8059894
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 462)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

Shogun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 MEDLINE 10737800
 PUBMED
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=QV4-NN0039-290
 300-154-a01&t3=2000-03-25&t4=1)
 Seq primer: puc 18 forward

High quality sequence start: 22
 High quality sequence stop: 234.
 Location/Qualifiers
 1. .462

FEATURES

Location/Qualifiers
 1. .462
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="NN0039"
 /note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 62.6%; Score 19.4; DB 2; Length 462;
 Best Local Similarity 79.3%; Pred. No. 8.4e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 2 AGTGAAGCGGTGACGATATCAGCGGC 30
 ||||| ||||| ||||| ||||| |||||
 Db 415 AGTGAAGCGGTGTCATCATGCTGACTGC 387

RESULT 27

CN195909/c
 LOCUS
 DEFINITION TgESTzyl91a10.y1 TgVEG118 Tachyzoite cDNA Library-2 Toxoplasma
 gondii cDNA clone TgESTzyl91a10.y1 5' similar to SM:TBB_TOXGO
 P10878 TUBULIN BETA CHAIN. ;, mRNA sequence.
 CN195909
 CN195909.1 GI:46220848
 EST.
 SOURCE Toxoplasma gondii
 ORGANISM Toxoplasma gondii
 Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
 Sarcocystidae; Toxoplasma.
 1 (bases 1 to 479)
 Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,
 Clifton, S., Pape, D., Martin, J., Wyllie, T., Dante, M., Marra, M.,
 Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M.,
 Ritter, E., Bennett, J., Franklin, C., Tsagareishvili, R., Ronko, I.,
 Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.

Accession
 Version
 Keywords
 Source
 Organism

REFERENCE

1 (bases 1 to 479)
 Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,
 Clifton, S., Pape, D., Martin, J., Wyllie, T., Dante, M., Marra, M.,
 Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M.,
 Ritter, E., Bennett, J., Franklin, C., Tsagareishvili, R., Ronko, I.,
 Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.
 Toxoplasma EST Project
 Unpublished (2001)
 Contact: Clifton, S.

TITLE

Toxoplasma EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: toxo@watson.wustl.edu
 Contact David Sibley (toxost@borcim.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: -40UP from Gibco
 High quality sequence stop: 402.

FEATURES

source

Location/Qualifiers
 1. .479
 /organism="Toxoplasma gondii"
 /mol_type="mRNA"
 /strain="VEG"
 /db_xref="taxon:5811"
 /clone="TgESTzyl91a10.y1"
 /dev_stage="Tachyzoite"
 /lab_host="GC10"

/clone_lib="TgVEG118 Tachyzoite cDNA Library-2"
 /note="Vector: pBluescript SK; Site 1: EcoRI; Site 2:
 XhoI; The library was constructed by Keliang Tang, Robert
 Cole and L. David Sibley at Washington University. cDNAs
 were synthesized from poly(A)+ RNAbly oligo d(T) priming,


```

Query Match      62.6%; Score 19.4; DB 7; Length 534;
Best Local Similarity 79.3%; Pred. No. 8.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTCAAGCGGTGACAGCATATCAGACGGCT 31
    |||||
Db 151 GGGAGCGGAGACAGCAGGTTCACACGGCT 123

RESULT 30
CF247429/c
LOCUS
DEFINITION
TgESTzy167c07.y1 TgMAS Tachyzoite cDNA Library Toxoplasma gondii
cDNA clone TgESTzy167c07.y1 5' similar to SW:TBB_TOKGO P10878
TUBULIN BETA CHAIN. ; mRNA sequence.

ACCESSION
CF247429
VERSION
CF247429.1 GI:33480271
KEYWORDS
EST.
SOURCE
Toxoplasma gondii
ORGANISM
Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
REFERENCE
1 (bases 1 to 541)
AUTHORS
Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
Ritter,E., Bennett,J., Franklin,C., Teagareishvili,R., Ronko,I.,
Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
Toxoplasma EST Project
Unpublished (2001)
TITLE
Toxoplasma EST Project
JOURNAL
COMMENT
Contact: Clifton, S.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxowatson.wustl.edu
Contact David Sibley (toxoeat@borcim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco
High quality sequence stop: 541.

FEATURES
Location/Qualifiers
1..541
    /organism="Toxoplasma gondii"
    /mol_type="mRNA"
    /strain="Tachyzoite"
    /db_xref="taxon:5811"
    /clone="TgESTzy167c07.y1"
    /dev_stage="Tachyzoite"
    /lab_host="Electroten Blue cells (Stratagene)"
    /clone_lib="TgMAS Tachyzoite cDNA Library"
    /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI. The cDNA library was constructed by Keliang Tang,
and Robert Cole at Washington University. cDNA was
synthesized from Poly(A)+ mRNA using an oligo-d(T) primer
containing a XhoI site. Following second strand synthesis,
EcoRI adapters were ligated to the cDNA, and products were
size-selected on sephacryl S500. The cDNA were
directionally cloned into the EcoRI/XhoI prepared
pBluescript II SK+ vector, and electroporated into
Electroten Blue cells (Stratagene). The library may
contain a small percentage of host or bacterial
contaminants."

ORIGIN
Query Match      62.6%; Score 19.4; DB 7; Length 541;
Best Local Similarity 79.3%; Pred. No. 8.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTCAAGCGGTGACAGCATATCAGACGGCT 31
    |||||
Db 300 GGGAGCGGAGACAGCAGGTTCACACGGCT 272

RESULT 32
CF247429/c
LOCUS
DEFINITION
TgESTzy124d09.y1 TgMAS Tachyzoite cDNA Library Toxoplasma gondii
cDNA clone TgESTzy124d09.y1 5' similar to SW:TBB_TOKGO P10878
TUBULIN BETA CHAIN. ; mRNA sequence.

ACCESSION
CD217441
VERSION
CD217441.1 GI:30957414
KEYWORDS
EST.
SOURCE
Toxoplasma gondii
ORGANISM
Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
REFERENCE
1 (bases 1 to 545)
AUTHORS
Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
Ritter,E., Bennett,J., Franklin,C., Teagareishvili,R., Ronko,I.,
Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
Toxoplasma EST Project
Unpublished (2001)
TITLE
Toxoplasma EST Project
JOURNAL
COMMENT
Contact: Clifton, S.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxowatson.wustl.edu
Contact David Sibley (toxoeat@borcim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..545
    /organism="Toxoplasma gondii"
    /mol_type="mRNA"
    /strain="Tachyzoite"
    /db_xref="taxon:5811"
    /clone="TgESTzy124d09.y1"
    /dev_stage="Tachyzoite"
    /lab_host="Electroten Blue cells (Stratagene)"
    /clone_lib="TgMAS Tachyzoite cDNA Library"
    /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI. The cDNA library was constructed by Keliang Tang,
and Robert Cole at Washington University. cDNA was
synthesized from Poly(A)+ mRNA using an oligo-d(T) primer
containing a XhoI site. Following second strand synthesis,
EcoRI adapters were ligated to the cDNA, and products were
size-selected on sephacryl S500. The cDNA were
directionally cloned into the EcoRI/XhoI prepared
pBluescript II SK+ vector, and electroporated into
Electroten Blue cells (Stratagene). The library may
contain a small percentage of host or bacterial
contaminants."

ORIGIN
Query Match      62.6%; Score 19.4; DB 6; Length 545;
Best Local Similarity 79.3%; Pred. No. 8.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAGCGGTGACAGCATATCAGACGGCT 31
    |||||
Db 87 GGGAGCGGAGACAGCAGGTTCACACGGCT 59

RESULT 33
CD217441/c
LOCUS
DEFINITION
TgESTzy124d09.y1 TgMAS Tachyzoite cDNA Library Toxoplasma gondii
cDNA clone TgESTzy124d09.y1 5' similar to SW:TBB_TOKGO P10878
TUBULIN BETA CHAIN. ; mRNA sequence.

ACCESSION
CD217441
VERSION
CD217441.1 GI:30957414
KEYWORDS
EST.
SOURCE
Toxoplasma gondii
ORGANISM
Toxoplasma gondii
Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
Sarcocystidae; Toxoplasma.
REFERENCE
1 (bases 1 to 545)
AUTHORS
Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
Ritter,E., Bennett,J., Franklin,C., Teagareishvili,R., Ronko,I.,
Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
Toxoplasma EST Project
Unpublished (2001)
TITLE
Toxoplasma EST Project
JOURNAL
COMMENT
Contact: Clifton, S.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: toxowatson.wustl.edu
Contact David Sibley (toxoeat@borcim.wustl.edu) for further
information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..545
    /organism="Toxoplasma gondii"
    /mol_type="mRNA"
    /strain="Tachyzoite"
    /db_xref="taxon:5811"
    /clone="TgESTzy124d09.y1"
    /dev_stage="Tachyzoite"
    /lab_host="Electroten Blue cells (Stratagene)"
    /clone_lib="TgMAS Tachyzoite cDNA Library"
    /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI. The cDNA library was constructed by Keliang Tang,
and Robert Cole at Washington University. cDNA was
synthesized from Poly(A)+ mRNA using an oligo-d(T) primer
containing a XhoI site. Following second strand synthesis,
EcoRI adapters were ligated to the cDNA, and products were
size-selected on sephacryl S500. The cDNA were
directionally cloned into the EcoRI/XhoI prepared
pBluescript II SK+ vector, and electroporated into
Electroten Blue cells (Stratagene). The library may
contain a small percentage of host or bacterial
contaminants."

ORIGIN
Query Match      62.6%; Score 19.4; DB 6; Length 545;
Best Local Similarity 79.3%; Pred. No. 8.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAGCGGTGACAGCATATCAGACGGCT 31
    |||||
Db 87 GGGAGCGGAGACAGCAGGTTCACACGGCT 59

RESULT 32
BM176352/c
LOCUS
DEFINITION
TgESTzyb34e03.y1 TgrH Tachyzoite Subtracted cDNA Library Toxoplasma
gondii cDNA clone TgESTzyb34e03.y1 5' similar to SW:TBB_TOKGO
P10878 TUBULIN BETA CHAIN. ; mRNA sequence.
BM176352
ACCESSION
BM176352.1 GI:17394570
VERSION

```

```

KEYWORDS
SOURCE      Tokoplasma gondii
ORGANISM    Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.

REFERENCE
1 (bases 1 to 546)
AUTHORS     Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
            Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
            Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
            Ritter,E., Bennett,J., Franklin,C., Tsagarishvili,R., Ronko,I.,
            Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
            Toxoplasma EST Project
            Unpublished (2001)
            Contact: Clifton, S.
            Toxoplasma EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: toxo@watson.wustl.edu
            Contact David Sibley (toxost@borcim.wustl.edu) for further
            information relating to organism, libraries, or clone availability.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 420.

TITLE
JOURNAL
COMMENT

FEATURES
source
1..546
    /organism="Toxoplasma gondii"
    /mol_type="mRNA"
    /strain="RH (Type I)"
    /db_xref="taxon:5811"
    /clone="TGESTzyb38f12.y1"
    /dev_stage="Tachyzoite"
    /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
    /clone_lib="TGRH Tachyzoite Subtracted cDNA Library"
    /note="Vector: pBluscript SK-; Site 1: EcoRI; Site 2:
XhoI; Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
originally constructed by K.L.Wan, Cambridge University.
cDNAs were synthesized from polyA RNAs by oligo d(T)
priming and directionally cloned into the EcoRI to XhoI
sites of the Lambda ZapII vector using the ZAP-cDNA
synthesis kit (Stratagene). The primary cDNA library was
mass excised as phagemid DNA using ExAssist helper phage
(Stratagene). Phagemid DNA was extracted by
phenol-chloroform method, and hybridized against a pool of
over-represented ESTs (N=12, from 5596 previous reads).
The subtracted library was electroporated into DH10B
(GeneHog, Invitrogen, Inc). WARNING: the library contains
a small percentage of cDNAs derived from the human host
cells. Library Source: David Sibley, Washington
University."

ORIGIN
Query Match      62.6%; Score 19.4; DB 4; Length 546;
Best Local Similarity 79.3%; Pred. No. 8.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGCGCT 31
    |||||
Db 44 GGGAGCGGAGACGACGTCACCGCT 16

RESULT 33
BM189407/c
LOCUS
DEFINITION
TgESTzyb38f12.y1 TGRH Tachyzoite Subtracted cDNA Library Toxoplasma
gondii cDNA clone TGESTzyb38f12.y1 5' similar to SW:TBH_TOXGO
BM189407
BM189407.1 GI:17525370
ACCESSION
VERSION
KEYWORDS
SOURCE      Tokoplasma gondii
ORGANISM    Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.

REFERENCE
1 (bases 1 to 546)
AUTHORS     Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajioke,J.A., White,M.,
            Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
            Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M.,
            Ritter,E., Bennett,J., Franklin,C., Tsagarishvili,R., Ronko,I.,
            Kennedy,S., Maguire,L., Waterston,R. and Wilson,R.
            Toxoplasma EST Project
            Unpublished (2001)
            Contact: Clifton, S.
            Toxoplasma EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: toxo@watson.wustl.edu
            Contact David Sibley (toxost@borcim.wustl.edu) for further
            information relating to organism, libraries, or clone availability.
            Seq primer: -40RP from Gibco
            High quality sequence stop: 420.

TITLE
JOURNAL
COMMENT

FEATURES
source
1..546
    /organism="Toxoplasma gondii"
    /mol_type="mRNA"
    /strain="RH (Type I)"
    /db_xref="taxon:5811"
    /clone="TGESTzyb38f12.y1"
    /dev_stage="Tachyzoite"
    /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
    /clone_lib="TGRH Tachyzoite Subtracted cDNA Library"
    /note="Vector: pBluscript SK-; Site 1: EcoRI; Site 2:
XhoI; Toxoplasma RH strain tachyzoites were grown in human
foreskin fibroblast cultures in vitro. The library was
originally constructed by K.L.Wan, Cambridge University.
cDNAs were synthesized from polyA RNAs by oligo d(T)
priming and directionally cloned into the EcoRI to XhoI
sites of the Lambda ZapII vector using the ZAP-cDNA
synthesis kit (Stratagene). The primary cDNA library was
mass excised as phagemid DNA using ExAssist helper phage
(Stratagene). Phagemid DNA was extracted by
phenol-chloroform method, and hybridized against a pool of
over-represented ESTs (N=12, from 5596 previous reads).
The subtracted library was electroporated into DH10B
(GeneHog, Invitrogen, Inc). WARNING: the library contains
a small percentage of cDNAs derived from the human host
cells. Library Source: David Sibley, Washington
University."

ORIGIN
Query Match      62.6%; Score 19.4; DB 4; Length 546;
Best Local Similarity 79.3%; Pred. No. 8.6e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACGACATATCAGCGCT 31
    |||||
Db 44 GGGAGCGGAGACGACGTCACCGCT 16

RESULT 34
BM189147/c
LOCUS
DEFINITION
TgESTzyb43b10.y1 TGRH Tachyzoite Subtracted cDNA Library Toxoplasma
gondii cDNA clone TGESTzyb43b10.y1 5' similar to SW:TBH_TOXGO
BM189147
BM189147.1 GI:17525102
ACCESSION
VERSION
KEYWORDS
SOURCE      Tokoplasma gondii
ORGANISM    Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
            Sarcocystidae; Toxoplasma.

```


Fax: 314 286 1810
 Email: toxo@watson.wustl.edu
 Contact David Sibley (toxost@borcim.wustl.edu) for further
 information relating to organism, libraries, or clone availability.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 424.
 Location/Qualifiers

FEATURES

source

1. .567
 /organism="Toxoplasma gondii"
 /mol_type="mRNA"
 /db_xref="taxon:5811"
 /clone="TgESTzf39g02.y1"
 /dev_stage="Tachyzoite"
 /lab_host="E.coli XL10"
 /clone_lib="TgME49 B7 Tachyzoite cDNA Library 2"
 /notes="Vector: pBluescript II SK+; Site 1: XhoI; Site 2:
 EcoRI; The cDNA library was constructed by Kelian Tang,
 and Robert Cole at Washington University. cDNA was
 synthesized from poly mRNA using an oligo-dT primer
 containing a XhoI site. Following second strand synthesis,
 EcoRI adapters were ligated to the cDNA, and products were
 size-selected on sephacryl S500. The cDNA were
 directionally cloned into the EcoRI/XhoI prepared
 pBluescript II SK+ vector, and transformed into XL10 cells
 (Stratagene). The library may contain a small percentage
 of host or bacterial contaminants."

ORIGIN

Query Match 62.6%; Score 19.4; DB 6; Length 567;
 Best Local Similarity 79.3%; Pred. No. 8.6e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGCGGCT 31
 | ||||| ||||| ||||| ||||| |||||
 Db 206 GCGAAGCGGAGACAGCGGTCACCGCT 178

RESULT 37

CB187210/c

LOCUS

CB187210

DEFINITION

gondii cDNA clone TgESTzf39g02.y1 5' similar to SW:TB_B_TOXGO

10878 TUBULIN BETA CHAIN. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

/strain="RH (Type I)"
 /db_xref="taxon:5811"
 /clone="TgESTzf39g02.y1"
 /dev_stage="Tachyzoite"
 /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
 /clone_lib="TgRH Tachyzoite Norm 5 cDNA Library"
 /notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
 XhoI; Toxoplasma RH strain tachyzoites were grown in human
 foreskin fibroblast cultures in vitro. The library was
 originally constructed by K.L.Wan, Cambridge University.
 cDNAs were synthesized from polyA RNAs by oligo d(T)
 priming and directionally cloned into the EcoRI to XhoI
 sites of the Lambda ZapII vector using the ZAP-cDNA
 synthesis kit (Stratagene). The primary cDNA library was
 mass excised as phagemid using ExAssist helper phage
 (Stratagene). Phagemid DNA was extracted by
 phenol-chloroform method, and hybridized against a pool of
 highly abundant genes which were derived from short-cycle
 PCR of the primary cDNA library. The normalized library
 was electroporated into DH10B (GeneHog, Invitrogen, Inc).
 WARNING: The library contains a small percentage of cDNAs
 derived from the human host cells."

ORIGIN

Query Match 62.6%; Score 19.4; DB 6; Length 574;
 Best Local Similarity 79.3%; Pred. No. 8.6e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGCGGCT 31
 | ||||| ||||| ||||| ||||| |||||
 Db 90 GCGAAGCGGAGACAGCGGTCACCGCT 62

RESULT 38

CB187263/c

LOCUS

CB187263

DEFINITION

gondii cDNA clone TgESTzf39g02.y1 5' similar to SW:TB_B_TOXGO

10878 TUBULIN BETA CHAIN. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CB187263
 TgESTzf39g02.y1 TgMAS Tachyzoite cDNA Library Toxoplasma gondii
 cDNA clone TgESTzf39g02.y1 5' similar to SW:TB_B_TOXGO P10878
 TUBULIN BETA CHAIN. ; mRNA sequence.

CB187263
 EST.

CB187263.1
 GI:47128821

Toxoplasma gondii

Toxoplasma gondii

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Sarcocystidae; Toxoplasma.

1 (bases 1 to 576)

Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioke, J.A., White, M.,

Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M.,

Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M.,

Ritter, E., Bennett, J., Franklin, C., Tsagareishvili, R., Ronko, I.,

Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.

Toxoplasma EST Project

Unpublished (2001)

Contact: Clifton, S.

Toxoplasma EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxo@watson.wustl.edu

Contact David Sibley (toxost@borcim.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco

High quality sequence stop: 497.

Location/Qualifiers

1. .576

/organism="Toxoplasma gondii"

/mol_type="mRNA"

/strain="Tachyzoite"

/db_xref="taxon:5811"

/clone="TgESTzf39g02.y1"

/dev_stage="Tachyzoite"

/lab_host="ElectroTen Blue cells (Stratagene)"
 /clone_lib="TgMAS Tachyzoite cDNA Library"
 /note="Vector: pBluescript II SK-; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed by Kelian Tang, and Robert Cole at Washington University. cDNA was synthesized from Poly(A)+ mRNA using an oligo-d(T) primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA, and products were size-selected on sephacryl S500. The cDNA were directionally cloned into the EcoRI/XhoI prepared pBluescript II SK+ vector, and electroporated into ElectroTen Blue cells (Stratagene). The library may contain a small percentage of host or bacterial contaminants."

ORIGIN

Query Match 62.6%; Score 19.4; DB 7; Length 576;
 Best Local Similarity 79.3%; Pred. No. 8.6e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGCGGCT 31
 | | | | | | | | | | | | | | | | | | | | |
 Db 86 GGGAGCGGAGACAGCAGCGGTACACCGCT 58

RESULT 39

CB187908/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB187908

TgESTzf49f03.y1 TgrH Tachyzoite Norm 7 cDNA Library Toxoplasma gondii

EST.

CB187908.1 GI:28206939

Toxoplasma gondii

Toxoplasma gondii

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Sarcocystidae; Toxoplasma.

1 (bases 1 to 578)

Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioka, J.A., White, M., Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.

Toxoplasma EST Project

Unpublished (2001)

Contact: Clifton, S.

Toxoplasma EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxo@watson.wustl.edu

Contact David Sibley (toxoeat@borcim.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40UP from Gibco

High quality sequence stop: 426.

Location/Qualifiers

1. .578

/organism="Toxoplasma gondii"

/mol_type="mRNA"

/strain="RH (Type I)"

/db_xref="taxon:5811"

/clone="TgESTzf49f03.y1"

/dev_stage="Tachyzoite"

/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"

/clone_lib="TgrH Tachyzoite Norm 7 cDNA Library"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Toxoplasma RH strain tachyzoites were grown in human foreskin fibroblast cultures in vitro. The library was originally constructed by K.L.Wan, Cambridge University. cDNAs were synthesized from polyA RNAs by oligo d(T) priming and directionally cloned into the EcoRI to XhoI

sites of the Lambda ZapII vector using the Zap-cDNA synthesis kit (Stratagene). The primary cDNA library was mass excised as phagemid using ExAssist helper phage (Stratagene). Phagemid DNA was extracted by phenol-chloroform method, and hybridized against a pool of highly abundant genes which were derived from short-cycle PCR of the primary cDNA library. The normalized library was electroporated into DH10B (GeneHog, Invitrogen, Inc.). WARNING: the library contains a small percentage of cDNAs derived from the human host cells."

ORIGIN

Query Match 62.6%; Score 19.4; DB 6; Length 578;
 Best Local Similarity 79.3%; Pred. No. 8.6e+02;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 GTGAAGCGGTGACAGCATATCAGCGGCT 31
 | | | | | | | | | | | | | | | | | | | | |
 Db 409 GGGAGCGGAGACAGCAGCGGTACACCGCT 381

RESULT 40

CB382287/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CB382287

TgESTzyg76d01.y1 TgME49 3 day invitro bradyzoite Toxoplasma gondii

EST.

CB382287.1 GI:29131583

Toxoplasma gondii

Toxoplasma gondii

Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

Sarcocystidae; Toxoplasma.

1 (bases 1 to 585)

Tang, K., Cole, R., Fogarty, S., Sibley, L.D., Ajioka, J.A., White, M., Clifton, S., Pape, D., Martin, J., Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Waterston, R. and Wilson, R.

Toxoplasma EST Project

Unpublished (2001)

Other ESTs: zyg76d01.x1

Contact: Clifton, S.

Toxoplasma EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: toxo@watson.wustl.edu

Contact David Sibley (toxoeat@borcim.wustl.edu) for further

information relating to organism, libraries, or clone availability.

Seq primer: -40RP from Gibco

High quality sequence stop: 489.

Location/Qualifiers

1. .585

/organism="Toxoplasma gondii"

/mol_type="mRNA"

/db_xref="taxon:5811"

/clone="TgESTzyg76d01.y1"

/dev_stage="3 day"

/lab_host="ElectroTen Blue cells"

/clone_lib="TgME49 3 day invitro bradyzoite"

/note="Vector: pBluescript SK; Site 1: NotI; Site 2: SmaI; cDNA was reverse transcribed using Superscript II and an anchored (7wobble) oligo dT primer, containing a NOTI restriction site, according to standard protocols. Blunt second strand cDNA was digested with NOTI restriction endonuclease, purified and size fractionated with a Siesep 400 spun column (Amersham) and directionally cloned into the SmaI and NotI sites of pBluescript SK. Following electroporation into ElectroTen Blue cells (Stratagene), 5 x 10⁶ primary CFU were bottle amplified in

semisolid culture media [2xLB+1% Seaprep agarosed (BMA)]. After 48 hrs at 30 degrees Celcius, cells were harvested, pooled and stored in 2xLB + 1% glycerol."

ORIGIN

```

Query Match      62.6%; Score 19.4; DB 6; Length 585;
Best Local Similarity 79.3%; Pred. No. 8.7e+02;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 GTGAAGCGGTGACAGCATATCAGACGGCT 31
Db 416 GGGAAAGCGGAGACAGCAGGTTCACACCGCT 388

```

Search completed: June 4, 2005, 11:46:50
Job time : 1896.65 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 04:22:34 ; Search time 159.178 Seconds
(without alignments)
780.981 Million cell updates/sec

Title: US-09-674-277-27
Perfect score: 21
Sequence: 1 gtagataggcaacaatga 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	100.0	21	3	Aaz36127 Primer de
2	21	100.0	1181	3	Aaz36102 Nucleic a
3	18	85.7	5402	4	Abi14236 Drosophil
4	17.4	82.9	1139	12	Adm72155 Human NTR
5	17.4	82.9	2643	11	Adm02069 Human CDN
6	16.8	80.0	240	6	Abk77060 Bacillus
7	16.8	80.0	2566	2	Aaz27256 Human sec
8	16.8	80.0	2566	8	Ada39862 Human sec
9	16.8	80.0	2566	9	Adb91161 Human sec
10	16.8	80.0	3883	10	Adc71228 Novel hum
11	16.8	80.0	83709	12	Adg19964 Human sof
12	16.8	80.0	110000	4	Continuation (14 o
13	16.8	80.0	110000	6	Continuation (14 o
14	16.8	80.0	110000	6	Continuation (14 o
15	16.8	80.0	110000	6	Continuation (14 o
16	16.8	80.0	110000	12	Continuation (14 o
17	16.4	78.1	9091	13	Adt05522 Haemophil
18	16.4	78.1	98800	12	Continuation (4 of
19	16.4	78.1	98800	13	Continuation (4 of
20	16.4	78.1	110000	2	Ata42063 Haemophil

C	21	16.4	78.1	254366	8	ABZ23704
	22	16.4	78.1	308766	13	ADT05738
	23	16.2	77.1	150	2	AAI12660
C	24	16.2	77.1	294	10	ABX87284
	25	16.2	77.1	602	13	ACN52890
C	26	16.2	77.1	624	13	ACN53918
	27	16.2	77.1	1964	13	ADS54785
C	28	16.2	77.1	2210	12	ADQ63615
	29	16.2	77.1	4045	4	ABU23258
	30	16.2	77.1	4576	10	ADD48369
	31	16.2	77.1	4576	10	ADE55806
C	32	16.2	77.1	44848	3	AAA75080
	33	16.2	77.1	44848	10	ADG88832
C	34	16.2	77.1	44848	12	ADL16411
	35	16.2	77.1	44848	12	ADM48748
C	36	16.2	77.1	55827	8	ACA60949
	37	16.2	77.1	58337	13	ADG13671
C	38	16.2	77.1	58337	13	ADS36454
	39	16.2	77.1	64423	13	ADS36462
C	40	16.2	77.1	91000	9	AAU61326
	41	16.2	77.1	96596	10	ADE95968
	42	16.2	77.1	96597	9	ADA02720
	43	16.2	77.1	96597	10	ADB72458
C	44	16.2	77.1	110000	13	ABD32627.3
	45	16.2	77.1	133955	11	ACN45170

ALIGNMENTS

RESULT 1
AAZ36127
ID AAZ36127 standard; DNA; 21 BP.
XX
AC AAZ36127;
DT 11-FEB-2000 (first entry)
DE Primer derived from a nucleic acid sequence specific to EHEC.
XX Enterohemorrhagic Escherichia coli; EHEC; katP gene; E. coli O157:H7;
KW IS91; virulence factor; enterohemolysine; ehly; intimin; eae; virK gene;
KW PCR primer; probe; ss.
XX Synthetic.
OS Escherichia coli.
XX WO9955908-A2.
XX PD 04-NOV-1999.
XX 27-APR-1999; 99WO-FR001000.
XX 28-APR-1998; 98FR-00005329.
PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.
PI Frechon DTM, Laure FC, Thierry D;
XX WPI; 2000-013443/01.
XX New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.
XX Claim 5; Page 27; 48pp; French.
CC AAZ36103-27 represent fragments derived from nucleic acid sequences
CC specific to enterohemorrhagic Escherichia coli (EHEC). The fragments are
CC derived from two sequences. The first (AAZ36101) is 99.9% homologous to
CC the katP gene of E. coli O157:H7 (nucleotides 407-1489 of AAZ36101), and
CC 95.8% homologous with IS91 of E. coli (nucleotides 1-406 of AAZ36102).
CC The second sequence (AAZ36102) is associated with the presence of

CC virulence factors enterohemolysine (ehly) and intimin (eae). Nucleotides
CC 237-570 of AAZ36102 also have 68% homology with the virK gene which codes
CC for virulence proteins of Shigella flexneri. Both sequences are of
CC plasmid origin. The fragments are used as PCR primers and probes for the
CC detection of E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC),
CC in human or animal samples, foods or the environment. The fragments are
CC also useful for epidemiological studies

XX
SQ Sequence 21 BP; 9 A; 3 C; 6 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 3; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
|||||
Db 1 GTGAGATAGGCACACAAATGA 21

RESULT 2

AAZ36102
ID AAZ36102 standard; DNA; 1181 BP.

XX AC AAZ36102;

XX DT 11-FEB-2000 (first entry)

XX Nucleic acid sequence specific to enterohemorrhagic Escherichia coli.

XX Enterohemorrhagic Escherichia coli; EHEC; virulence factor;

XX enterohemolysine; ehly; intimin; eae; virK gene; E. coli O157:H7; ds.

XX Escherichia coli.

XX WO9955908-A2.

XX PD 04-NOV-1999.

XX PF 27-APR-1999; 99WO-FR001000.

XX PR 28-APR-1998; 98FR-00005329.

XX PA (SNFI) PASTEUR SANOFI DIAGNOSTICS.

XX PI Frechon DTM, Laure FC, Thierry D;

XX WPI; 2000-013443/01.

XX New nucleic acid containing sequences specific to enterohemorrhagic
PT Escherichia coli, particularly serotype O157:H7, used for detecting these
PT bacteria in food.

XX Claim 1; Fig 2; 48pp; French.

XX The present sequence is specific to enterohemorrhagic Escherichia coli
CC (EHEC). The sequence associated with the presence of virulence factors
CC enterohemolysine (ehly) and intimin (eae). Nucleotides 237-570 also have
CC 68% homology with the virK gene which codes for virulence proteins of
CC Shigella flexneri. The present sequence is of plasmid origin. Fragments
CC of the present sequence are used, as probes and primers, for detection of
CC E. coli O157:H7 and other enterohemorrhagic E. coli (EHEC), in human or
CC animal samples, foods or the environment. The fragments are also useful
CC for epidemiological studies

XX SQ Sequence 1181 BP; 305 A; 317 C; 277 G; 282 T; 0 U; 0 Other;

Query Match 100.0%; Score 21; DB 3; Length 1181;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
|||||

Db 883 GTGAGATAGGCACACAAATGA 903
|||||

RESULT 3

ABL14236

ID ABL14236 standard; cDNA; 5402 BP.

XX AC ABL14236;

XX DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 37190.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX DR P-PSDB; ABB70133.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.

XX Claim 1; SEQ ID NO 37190; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 5402 BP; 1296 A; 1265 C; 1152 G; 1689 T; 0 U; 0 Other;

Query Match 85.7%; Score 18; DB 4; Length 5402;

Best Local Similarity 100.0%; Pred. No. 65;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGATAGGCACACAAATGA 21
|||||

Db 3041 AGATAGGCACACAAATGA 3058

RESULT 4

ADM72155

ID ADM72155 standard; cDNA; 1139 BP.

XX AC ADM72155;

XX DT 03-JUN-2004 (first entry)

XX Human NTRAN polypeptide encoding cDNA (clone ID 6981317CB1).

XX NTRAN; neurotransmission-associated protein; cytostatic;
KW antiarteriosclerotic; anti-HIV; antiallergic; cerebroprotective;

KW antiparkinsonian; anticonvulsant; dermatological; neurotropic; neuroprotective; antiinflammatory; antithyroid; antiarthritic; anorectic;
KW cardiant; hypotensive; hepatotropic; gene therapy; human; gene; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
PH CDS 42..1025
FT /*tag= a
FT /product= "NTRAN"

XX WO2004022705-A2.

XX 18-MAR-2004.

XX 02-SEP-2003; 2003WO-US027411.

XX 04-SEP-2002; 2002US-0408383P.

XX 05-SEP-2002; 2002US-0408781P.

XX 26-SEP-2002; 2002US-0414221P.

XX 13-NOV-2002; 2002US-0426483P.

XX 06-DEC-2002; 2002US-0431566P.

XX 17-DEC-2002; 2002US-0434317P.

XX 03-JAN-2003; 2003US-0437763P.

XX (INCY-) INCYTE CORP.

XX Lee SY, Elliott VS, Hafalia AJA, Burford N, Sprague WW;
PI Griffin JA, Yang YG, Chawla NK, Baughn MR, Becha SD, Khare R;
PI Thornton MB, Mason PM, Gietzen KJ, Ison CH, Marquis JP, Swarnakar A;
PI Ramkumar J, Jin P, Richardson TW, Tran UK;
XX

XX WPI; 2004-269571/25.

XX P-PSDB; ADM72125.

XX New human neurotransmission-associated proteins and polynucleotides for
PT diagnosing, preventing or treating diseases or conditions associated with
PT aberrant protein expression, e.g. cancer, hepatitis, AIDS, obesity or
PT stroke.

XX Claim 5; SEQ ID NO 36; 216pp; English.

XX The invention relates to human neurotransmission-associated proteins
CC (NTRAN) and encoding polynucleotides. The NTRAN polypeptides can be
CC expressed by standard recombinant methodology. The polypeptides and
CC polynucleotides are useful in diagnosing, preventing or treating diseases
CC or conditions associated with the decreased expression or overexpression
CC of NTRAN, such as autoimmune/inflammatory, cardiovascular, neurological,
CC developmental, cell proliferative, transport, psychiatric, metabolic or
CC endocrine disorders. These diseases may include AIDS, allergies, atopic
CC dermatitis, arthritis, thyroiditis, obesity, Parkinson's disease,
CC Alzheimer's disease, stroke, epilepsy, myocardial infarction,
CC hypertension, cancer, atherosclerosis or hepatitis. These are also useful
CC in assessing the effects of exogenous compounds on the expression of
CC nucleic acid and amino acid sequences of NTRAN. The NTRAN or its
CC fragments are useful in screening compounds for effectiveness as agonist
CC or antagonist of the polypeptides, or in altering the expression of the
CC target polynucleotide and compounds that specifically bind to or modulate
CC the activity of the polypeptide. The microarray is useful in monitoring
CC or measuring protein-protein interactions, drug-target interactions, and
CC gene expression profiles. The present sequence represents a specific
CC example of a polynucleotide encoding a human NTRAN polypeptide.

XX Sequence 1139 BP; 320 A; 269 C; 271 G; 279 T; 0 U; 0 Other;

Query Match 82.9%; Score 17.4; DB 12; Length 1139;
Best Local Similarity 94.7%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAGATAGGCACACAAATGA 21

Db 27 GAGATAGGCACACAAATGA 45

RESULT 5
ADM02069
ID ADM02069 standard; cDNA; 2643 BP.
XX
XX ADM02069;
XX
XX 20-MAY-2004 (first entry)
XX
XX Human cDNA of the invention SEQ ID NO:754.
XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX
XX Homo sapiens.
XX EP1347046-A1.
XX
XX 24-SEP-2003.
XX
XX 12-APR-2002; 2002EP-00008400.
XX
XX 22-MAR-2002; 2002JP-00137785.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
XX

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI; 2003-723558/69.
XX P-PSDB; ADM04512.

XX New polynucleotides and polypeptides are useful in gene therapy, for
PT developing a diagnostic marker or medicines for regulating their
PT expression and activity, or as a target of gene therapy.

XX Claim 1; SEQ ID NO 754; 305pp; English.

XX The invention relates to a novel human polynucleotide and the encoded
CC polypeptide. A polynucleotide of the invention may have a use in gene
CC therapy. An oligonucleotide of the invention ADM060773 is useful
CC as a primer for synthesizing the polynucleotide or as a probe for
CC detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC useful in gene therapy, for developing a diagnostic marker or medicines
CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.

XX Sequence 2643 BP; 648 A; 674 C; 743 G; 578 T; 0 U; 0 Other;

Query Match 82.9%; Score 17.4; DB 11; Length 2643;
Best Local Similarity 94.7%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAGATAGGCACACAAATGA 21

Db 1325 GAGATAGGCACACAAATGA 1343

RESULT 6

ABK77060

ID ABK77060 standard; DNA; 240 BP.

XX
XX ABK77060;

XX
XX 13-AUG-2002 (first entry)

XX Bacillus licheniformis genomic sequence tag (GST) #4351.

XX Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.

XX Bacillus licheniformis.
 XX WO200229113-A2.
 XX
 PD 11-APR-2002.
 XX
 XX 05-OCT-2001; 2001WO-US031437.
 XX
 XX 06-OCT-2000; 2000US-00680598.
 PR 27-MAR-2001; 2001US-0279526P.
 XX
 PA (NOVO) NOVOZYMES BIOTECH INC.
 PA (NOVO) NOVOZYMES AS.
 XX
 PI Berka R, Clausen IG;
 XX
 DR WPI; 2002-416684/44.
 XX
 XX Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second Bacillus
 PT cells, by using substrate containing Bacillus genomic sequenced tag
 PT array.
 XX
 PS Claim 4; SEQ ID NO 4351; 200pp; English.
 XX
 CC The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive follow
 CC -up characterisation is unnecessary, when one spot on an array equals one
 CC gene or one open reading frame, since sequence information is available.
 CC This sequence represents a genomic sequence tag (GST) used in the method
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 240 BP; 69 A; 50 C; 59 G; 62 T; 0 U; 0 Other;
 Query Match 80.0%; Score 16.8; DB 6; Length 240;
 Best Local Similarity 90.0%; Pred. No. 1.7e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TGAGATAGGCACACAAATGA 21
 ||||| ||||| ||||| ||||| |||||
 Db 60 TGAGATGGGCACACAACTTGA 79
 RESULT 7
 AAZ27256/c
 ID AAZ27256 standard; cDNA; 2566 BP.
 XX
 XX AAZ27256;
 AC
 DT 23-NOV-1999 (first entry)
 XX
 DE Human secreted protein cDNA encoding gene 24.
 XX
 KW Secreted protein; human; gene therapy; diagnosis; treatment; cancer;
 KW protein therapy; tumor; neurodegenerative disorder; blood disorder; AIDS;
 KW developmental abnormality; leukemia; immune system; autoimmune disease;
 KW hepatic disease; renal disease; inflammation; allergy; schizophrenia;

KW Alzheimer's disease; cognitive disorder; arthritis; infection; psoriasis;
 KW transplant rejection; diabetes; asthma; sepsis; acne; metabolic disorder;
 KW cardiovascular disorder; food additive; preservative; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO9946289-A1.
 XX
 PD 16-SEP-1999.
 XX
 PF 11-MAR-1999; 99WO-US005721.
 XX
 PR 12-MAR-1998; 98US-0077686P.
 PR 12-MAR-1998; 98US-0077687P.
 PR 12-MAR-1998; 98US-0077696P.
 PR 12-MAR-1998; 98US-0077714P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 XX Ruben SM, Ferrie AM, Rosen CA, Florence C, Young PE, Yu G, Ni J;
 PI
 XX WPI; 1999-551363/46.
 XX
 DR P-PSDB; AAY45280, AAY45365.
 DR
 XX New isolated human genes, useful for diagnosis and treatment of, e.g.
 PT cancers.
 PT
 PS Claim 1a; 165-166; 306pp; English.
 XX
 CC This invention describes novel human genes and the secreted proteins they
 CC encode. The polynucleotides and their corresponding secreted polypeptides
 CC are useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. Also pathological conditions can be
 CC diagnosed by determining the amount of the new polypeptides in a sample
 CC or by determining the presence of mutations in the new polynucleotides.
 CC Specific uses are described for each of the polynucleotides of the
 CC invention, based on which tissues they are most highly expressed in, and
 CC include developing products for the diagnosis or treatment of cancer,
 CC tumors, neurodegenerative disorders, developmental abnormalities, blood
 CC disorders, leukemias, diseases of the immune system, autoimmune diseases,
 CC hepatic and renal disease, inflammation, allergies, Alzheimer's and
 CC cognitive disorders, schizophrenia, arthritis, infections, AIDS,
 CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
 CC cardiovascular disorders, and metabolic disorders. The polypeptides or
 CC polynucleotides can also be used as food additives or preservatives. The
 CC polypeptides are also useful for identifying their binding partners.
 CC AAZ27233-27265 encode human secreted proteins described in the method of
 CC the invention
 XX
 SQ Sequence 2566 BP; 802 A; 491 C; 466 G; 791 T; 0 U; 16 Other;
 Query Match 80.0%; Score 16.8; DB 2; Length 2566;
 Best Local Similarity 90.0%; Pred. No. 2.4e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TGAGATAGGCACACAAATGA 21
 ||||| ||||| ||||| ||||| |||||
 Db 1253 TGAAGAAAGGCACACAAATGA 1234
 RESULT 8
 ADA39862/c
 ID ADA39862 standard; cDNA; 2566 BP.
 XX
 XX ADA39862;
 AC
 XX 20-NOV-2003 (first entry)
 DT
 XX Human secreted protein encoding cDNA.
 DE
 XX Human; secreted protein; cancer; hyperproliferative disorder;
 KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
 KW anaemia; allergic reaction; asthma; cardiovascular disorder;

KW wound healing; cytostatic; immunosuppressive; neurotropic; neuroprotective;
KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
KW vulvular; cardiac; gene therapy; ss.

XX Homo sapiens.

XX WO2002102993-A2.

XX 27-DEC-2002.

XX 19-MAR-2002; 2002WO-US008123.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-175238/17.

XX New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating cancer or other hyperproliferative disorder,
PT asthma, allergies or AIDS.

XX Claim 9; SEQ ID NO 244; 3205pp; English.

XX The invention relates to novel genes ADA39629-ADA40565 and proteins
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC fragments, and agonists or antagonists that bind to the polypeptide are
CC useful for preparing a diagnostic or pharmaceutical composition for
CC diagnosing or treating cancer or other hyperproliferative disorder. The
CC polypeptides and nucleic acid molecules are also useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC or other hyperproliferative disorders including neoplasms, autoimmune
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,
CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunological probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2566 BP; 802 A; 491 C; 466 G; 791 T; 0 U; 16 Other;

Query Match 80.0%; Score 16.8; DB 8; Length 2566;
Best Local Similarity 90.0%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGAGATAGGCACCAACATGA 21

Db 1253 TGAAGGAGGCACCAACATGA 1234

RESULT 9

ID ADB91161/c

ADB91161 standard; cDNA; 2566 BP.

XX

AC ADB91161;

XX 04-DEC-2003 (first entry)

XX Human secreted protein cDNA #SEQ ID 107.

XX Secreted protein; gene therapy; antidiabetic; diabetes; human; ss.

XX Homo sapiens.

XX WO2003004622-A2.

XX 16-JAN-2003.

XX 19-MAR-2002; 2002WO-US008124.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-229407/22.

XX Nucleic acid encoding a human secreted protein is useful in diagnosing or

XX treating diabetes or conditions related to diabetes.

XX Claim 9; SEQ ID NO 107; 1537pp; English.

XX The invention relates to isolated nucleic acid molecules ADB91065-
CC ADB91448 and ADB91835-ADB91911 encoding human secreted proteins ADB91449-
CC ADB91834. Also disclosed is a recombinant vector comprising a
CC polynucleotide of the invention, and a recombinant host cell comprising
CC the recombinant vector. The polypeptide of the invention is useful in
CC identifying a binding partner by contacting the polypeptide with a
CC binding partner, and determining whether the binding partner increases or
CC decreases activity of the polypeptide. The polypeptide, polynucleotide,
CC antibody or its fragment, agonist or antagonist are useful for preparing
CC a pharmaceutical composition for diagnosing or treating diabetes or
CC conditions related to diabetes. The present sequence is that of the human
CC immunoglobulin Fc portion used to generate fusion proteins, increasing
CC the stability of the fused protein as compared to the secreted protein
CC only. Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2566 BP; 802 A; 491 C; 466 G; 791 T; 0 U; 16 Other;

Query Match 80.0%; Score 16.8; DB 9; Length 2566;
Best Local Similarity 90.0%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGAGATAGGCACCAACATGA 21

Db 1253 TGAAGGAGGCACCAACATGA 1234

RESULT 10

ADB91228/c

ID ADB91228 standard; DNA; 3883 BP.

XX ADE71228;

XX 29-JAN-2004 (first entry)

XX Novel human protein coding sequence #44.

XX human; novel protein; drug; gene; ds.

XX Homo sapiens.

XX

PN JP2002345493-A.
XX
PD 03-DEC-2002.
XX
PF 29-MAR-2001; 2002JP-00049046.
XX
PR 29-MAR-2001; 2001JP-00095524.
XX
XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.
XX
XX WPI; 2003-460885/44.
DR P-PSDB; ADE71290.
XX
XX A gene and a protein encoded by it, used in drugs.
PT
PS Claim 1; SEQ ID NO 44; 257pp; Japanese.
XX
XX The invention comprises the amino acid and coding sequences of novel
CC human proteins. The DNA and protein sequences of the invention are used
CC in drugs. The present DNA sequence encodes a novel human protein of the
CC invention.
XX
SQ Sequence 3883 BP; 1259 A; 650 C; 780 G; 1194 T; 0 U; 0 Other;
Query Match 80.0%; Score 16.8; DB 10; Length 3883;
Best Local Similarity 90.0%; Pred. No. 2.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 TGAGATAGGCACACAAATGA 21
DB 912 TGAAGGACACAAATGA 893
RESULT 11
ADQ19964
ID ADQ19964 standard; DNA; 83709 BP.
XX
AC ADQ19964;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 2784.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
XX 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 2784; 210pp; English.
PS
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the

CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytotstatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 83709 BP; 26383 A; 16046 C; 15839 G; 25441 T; 0 U; 0 Other;

Query Match 80.0%; Score 16.8; DB 12; Length 83709;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGAGATAGGCACACAAATGA 21
DB 77880 TGAGATAAGCACAGCAATGA 77899

RESULT 12

AAK95240_13
Continuation (14 of 15) of AAK95240 from base 1300001 (Human neuregulin-1 gene.)
WP Sequence split into 15 fragments LOCUS AAK95240 Accession Aak95240

WP	Fragment Name	Begin	End
WP	AAK95240_00	1	110000
WP	AAK95240_01	100001	210000
WP	AAK95240_02	200001	310000
WP	AAK95240_03	300001	410000
WP	AAK95240_04	400001	510000
WP	AAK95240_05	500001	610000
WP	AAK95240_06	600001	710000
WP	AAK95240_07	700001	810000
WP	AAK95240_08	800001	910000
WP	AAK95240_09	900001	1010000
WP	AAK95240_10	1000001	1110000
WP	AAK95240_11	1100001	1210000
WP	AAK95240_12	1200001	1310000
WP	AAK95240_13	1300001	1410000
WP	AAK95240_14	1400001	1503900

Query Match 80.0%; Score 16.8; DB 4; Length 110000;
Best Local Similarity 90.0%; Pred. No. 4.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAATG 20
DB 86719 GTGTGATAGGCACAACTG 86738

RESULT 13

AAK96733_13
Continuation (14 of 15) of AAK96733 from base 1300001 (Human neuregulin-1 gene.)
WP Sequence split into 15 fragments LOCUS AAK96733 Accession Aak96733

WP	Fragment Name	Begin	End
WP	AAK96733_00	1	110000
WP	AAK96733_01	100001	210000
WP	AAK96733_02	200001	310000
WP	AAK96733_03	300001	410000
WP	AAK96733_04	400001	510000
WP	AAK96733_05	500001	610000
WP	AAK96733_06	600001	710000
WP	AAK96733_07	700001	810000
WP	AAK96733_08	800001	910000
WP	AAK96733_09	900001	1010000
WP	AAK96733_10	1000001	1110000
WP	AAK96733_11	1100001	1210000
WP	AAK96733_12	1200001	1310000
WP	AAK96733_13	1300001	1410000
WP	AAK96733_14	1400001	1503900

Query Match 80.0%; Score 16.8; DB 4; Length 110000;

```
Best Local Similarity 90.0%; Pred. No. 4.1e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

QY 1 GTGATAGGACCAACAATG 20
|||||
Db 86719 GTGTGATAGGCACACACTG 86738

RESULT 14
ABT00010_13
Continuation (14 of 15) of ABT00010 from base 1300001 (Human neuregulin 1 gene.)
WP Sequence split into 15 fragments LOCUS ABT00010 Accession Abt00010
WP Fragment Name Begin End
WP ABT00010_00 1 110000
WP ABT00010_01 100001 210000
WP ABT00010_02 200001 310000
WP ABT00010_03 300001 410000
WP ABT00010_04 400001 510000
WP ABT00010_05 500001 610000
WP ABT00010_06 600001 710000
WP ABT00010_07 700001 810000
WP ABT00010_08 800001 910000
WP ABT00010_09 900001 1010000
WP ABT00010_10 1000001 1110000
WP ABT00010_11 1100001 1210000
WP ABT00010_12 1200001 1310000
WP ABT00010_13 1300001 1410000
WP ABT00010_14 1400001 1503841

Query Match 80.0%; Score 16.8; DB 6; Length 110000;
Best Local Similarity 90.0%; Pred. No. 4.1e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

QY 1 GTGATAGGACCAACAATG 20
|||||
Db 86660 GTGTGATAGGCACACACTG 86679

RESULT 15
ABT01503_13
Continuation (14 of 15) of ABT01503 from base 1300001 (Human neuregulin 1 gene.)
WP Sequence split into 15 fragments LOCUS ABT01503 Accession Abt01503
WP Fragment Name Begin End
WP ABT01503_00 1 110000
WP ABT01503_01 100001 210000
WP ABT01503_02 200001 310000
WP ABT01503_03 300001 410000
WP ABT01503_04 400001 510000
WP ABT01503_05 500001 610000
WP ABT01503_06 600001 710000
WP ABT01503_07 700001 810000
WP ABT01503_08 800001 910000
WP ABT01503_09 900001 1010000
WP ABT01503_10 1000001 1110000
WP ABT01503_11 1100001 1210000
WP ABT01503_12 1200001 1310000
WP ABT01503_13 1300001 1410000
WP ABT01503_14 1400001 1503841

Query Match 80.0%; Score 16.8; DB 6; Length 110000;
Best Local Similarity 90.0%; Pred. No. 4.1e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

QY 1 GTGATAGGACCAACAATG 20
|||||
Db 86660 GTGTGATAGGCACACACTG 86679

RESULT 16
ADH77486_13
Continuation (14 of 15) of ADH77486 from base 1300001 (Human neuregulin gene.)
WP Sequence split into 15 fragments LOCUS ADH77486 Accession Adh77486
WP Fragment Name Begin End
```

```
WP ADH77486_00 1 110000
WP ADH77486_01 100001 210000
WP ADH77486_02 200001 310000
WP ADH77486_03 300001 410000
WP ADH77486_04 400001 510000
WP ADH77486_05 500001 610000
WP ADH77486_06 600001 710000
WP ADH77486_07 700001 810000
WP ADH77486_08 800001 910000
WP ADH77486_09 900001 1010000
WP ADH77486_10 1000001 1110000
WP ADH77486_11 1100001 1210000
WP ADH77486_12 1200001 1310000
WP ADH77486_13 1300001 1410000
WP ADH77486_14 1400001 1503841

Query Match 80.0%; Score 16.8; DB 12; Length 110000;
Best Local Similarity 90.0%; Pred. No. 4.1e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 2;

QY 1 GTGATAGGACCAACAATG 20
|||||
Db 86660 GTGTGATAGGCACACACTG 86679

RESULT 17
ADT05522/c
ID ADT05522 standard; DNA; 9091 BP.
XX
AC ADT05522;
XX
DT 02-DEC-2004 (first entry)
XX
DE Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 558.
XX
KW middle ear bacterial infection; nasopharynx bacterial infection; ds;
KW contig.
XX
OS Haemophilus influenzae.
XX
PN WO2004078949-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-US007001.
XX
PR 06-MAR-2003; 2003US-0453134P.
XX
PA (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI Bakaletz LO, Munson RS, Dyer DW;
XX
DR WPI; 2004-662422/64.
XX
PT New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT useful for treating or preventing NTHi bacterial infections of the middle
PT ear and/or nasopharynx.
XX
PS Example 1; SEQ ID NO 558; 88pp; English.
XX
CC The invention comprises nucleotide sequences (genes) from the genome of a
CC nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC sequences of the invention are useful for treating or preventing NTHi
CC bacterial infections of the middle ear and/or nasopharynx. The present
CC nucleic acid represents an NTHi contig sequence of the invention.
XX
SQ Sequence 9091 BP; 2708 A; 1684 C; 1817 G; 2878 T; 0 U; 4 Other;

Query Match 78.1%; Score 16.4; DB 13; Length 9091;
Best Local Similarity 94.4%; Pred. No. 4.5e+02; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 1;

QY 3 GAGATAGGACCAACAATG 20
```

Db 6349 GAGATAGGCACCAATG 6332
|||||
RESULT 18
ADN06353_3/c
Continuation (4 of 4) of ADN06353 from base 300001 (Human FLAP genomic DNA SEQ ID NO:1.
WP Sequence split into 4 fragments LOCUS ADN06353 Accession Adn06353
Fragment Name Begin End
WP ADN06353_0 1 110000
WP ADN06353_1 100001 210000
WP ADN06353_2 200001 310000
WP ADN06353_3 300001 398800
Query Match 78.1%; Score 16.4; DB 12; Length 98800;
Best Local Similarity 94.4%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 AGATAGGCACCAATGA 21
|||
Db 52123 AGGTAGGCACCAATGA 52106
|||
RESULT 19
ADS94372_3/c
Continuation (4 of 4) of ADS94372 from base 300001 (Human 5-lipoxygenase activating prob
WP Sequence split into 4 fragments LOCUS ADS94372 Accession AdS94372
Fragment Name Begin End
WP ADS94372_0 1 110000
WP ADS94372_1 100001 210000
WP ADS94372_2 200001 310000
WP ADS94372_3 300001 398800
Query Match 78.1%; Score 16.4; DB 13; Length 98800;
Best Local Similarity 94.4%; Pred. No. 6.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 AGATAGGCACCAATGA 21
|||
Db 52123 AGGTAGGCACCAATGA 52106
|||
RESULT 20
AAT42063_00
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063
Fragment Name Begin End
WP AAT42063_00 1 110000
WP AAT42063_01 100001 210000
WP AAT42063_02 200001 310000
WP AAT42063_03 300001 410000
WP AAT42063_04 400001 510000
WP AAT42063_05 500001 610000
WP AAT42063_06 600001 710000
WP AAT42063_07 700001 810000
WP AAT42063_08 800001 910000
WP AAT42063_09 900001 1010000
WP AAT42063_10 1000001 1110000
WP AAT42063_11 1100001 1210000
WP AAT42063_12 1200001 1310000
WP AAT42063_13 1300001 1410000
WP AAT42063_14 1400001 1510000
WP AAT42063_15 1500001 1610000
WP AAT42063_16 1600001 1710000
WP AAT42063_17 1700001 1810000
WP AAT42063_18 1800001 1830121
ID AAT42063 standard; DNA; 1830121 BP.
XX AAT42063;
AC AAT42063;
XX
DT 14-SEP-1999 (first entry)
XX
DE Haemophilus influenzae complete genome sequence.
XX

KW Genome; bacterium; Haemophilus influenzae; computer readable medium;
KW expression modulating fragment; regulation; gene expression; vector;
KW organism; open reading frame; ORF; ds.
OS Haemophilus influenzae.
XX
PN WO9633276-A1.
XX
PD 24-OCT-1996.
XX
PF 22-APR-1996; 96WO-US005320.
XX
PR 21-APR-1995; 95US-00426787.
PR 07-JUN-1995; 95US-00476102.
PR 07-JUN-1995; 95US-00487429.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (UJJO) UNIV JOHNS HOPKINS.
XX
PI Fleischmann RD, Adams MD, White O, Smith HO, Venter JC;
XX
DR WPI; 1996-485782/48.
XX
XX Haemophilus influenzae Rd genome recorded on computer readable medium -
PT useful for identifying commercially important nucleic acid fragments by
PT homology searching.
XX
PS Claim 1; Page 77.2-77.1091; 1291pp; English.
XX
CC This sequence represents the complete genome sequence of the bacterium
CC Haemophilus influenzae strain Rd. The invention relates to a computer
CC readable medium (CRM) having recorded upon it the complete H.influenzae
CC nucleotide sequence (I), a representative fragment of (I) or a nucleotide
CC sequence at least 99% identical to (I). By providing the full-length
CC genomic sequence in a computer readable form, it is possible to identify
CC commercially important nucleic acid fragments and expression modulating
CC fragments (EMFs) of the Haemophilus genome. The EMFs can be used to
CC regulate the expression of a nucleic acid molecule. Vectors and altered
CC organisms comprising the predicted ORFs can be used to produce any of the
CC polypeptide fragments of the H. influenzae Rd genome
XX
SQ Sequence 1830121 BP; 567399A; 350615C; 347389G; 564036T; 0U; 6820ther;
Query Match 78.1%; Score 16.4; DB 2; Length 110000;
Best Local Similarity 94.4%; Pred. No. 6.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3 GAGATAGGCACCAATG 20
|||||
Db 72931 GAGATAGGCACCAATG 72948
|||||
RESULT 21
ABZ23704/c
ID ABZ23704 standard; cDNA; 254366 BP.
XX
AC ABZ23704;
XX
DT 21-MAR-2003 (first entry)
XX
DE Human phosphatase genomic DNA #SEQ ID 3.
XX
KW Human; phosphatase; enzyme; chromosome 12; colon adenocarcinoma;
KW placenta; gene therapy; single nucleotide polymorphism; SNP; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(36..37,A)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
FT /replace(1414,M)
FT /*tag= b

FT /standard_name= "single nucleotide polymorphism"
FT replace(1743,C)
FT /tag= c
FT /standard_name= "single nucleotide polymorphism"
FT replace(2766,A)
FT /tag= d
FT /standard_name= "single nucleotide polymorphism"
FT replace(3155,T)
FT /tag= e
FT /standard_name= "single nucleotide polymorphism"
FT replace(5816,A)
FT /tag= f
FT /standard_name= "single nucleotide polymorphism"
FT replace(6074,T)
FT /tag= g
FT /standard_name= "single nucleotide polymorphism"
FT replace(9550,C)
FT /tag= h
FT /standard_name= "single nucleotide polymorphism"
FT replace(9644,G)
FT /tag= i
FT /standard_name= "single nucleotide polymorphism"
FT replace(16630,G)
FT /tag= j
FT /standard_name= "single nucleotide polymorphism"
FT replace(17957,A)
FT /tag= k
FT /standard_name= "single nucleotide polymorphism"
FT replace(18299,G)
FT /tag= l
FT /standard_name= "single nucleotide polymorphism"
FT replace(23521,G)
FT /tag= m
FT /standard_name= "single nucleotide polymorphism"
FT replace(28463,A)
FT /tag= n
FT /standard_name= "single nucleotide polymorphism"
FT replace(35221,T)
FT /tag= o
FT /standard_name= "single nucleotide polymorphism"
FT replace(41813,G)
FT /tag= p
FT /standard_name= "single nucleotide polymorphism"
FT replace(41957,T)
FT /tag= q
FT /standard_name= "single nucleotide polymorphism"
FT replace(42599,T)
FT /tag= r
FT /standard_name= "single nucleotide polymorphism"
FT replace(47819,A)
FT /tag= s
FT /standard_name= "single nucleotide polymorphism"
FT replace(51990, .51991,A)
FT /tag= t
FT /standard_name= "single nucleotide polymorphism"
FT replace(51992, .51993,T)
FT /tag= v
FT /standard_name= "single nucleotide polymorphism"
FT replace(51992,T)
FT /tag= u
FT /standard_name= "single nucleotide polymorphism"
FT replace(52788,A)
FT /tag= w
FT /standard_name= "single nucleotide polymorphism"
FT replace(59029,C)
FT /tag= x
FT /standard_name= "single nucleotide polymorphism"
FT replace(60776,T)
FT /tag= y
FT /standard_name= "single nucleotide polymorphism"
FT replace(61193,G)
FT /tag= z
FT /standard_name= "single nucleotide polymorphism"

FT variation
FT replace(62994,T)
FT /tag= aa
FT /standard_name= "single nucleotide polymorphism"
FT replace(63244,T)
FT /tag= ab
FT /standard_name= "single nucleotide polymorphism"
FT replace(65053,T)
FT /tag= ac
FT /standard_name= "single nucleotide polymorphism"
FT replace(68460,A)
FT /tag= ad
FT /standard_name= "single nucleotide polymorphism"
FT replace(69326,G)
FT /tag= ae
FT /standard_name= "single nucleotide polymorphism"
FT replace(73039,G)
FT /tag= af
FT /standard_name= "single nucleotide polymorphism"
FT replace(73084,G)
FT /tag= ag
FT /standard_name= "single nucleotide polymorphism"
FT replace(75205,A)
FT /tag= ah
FT /standard_name= "single nucleotide polymorphism"
FT replace(75491,C)
FT /tag= ai
FT /standard_name= "single nucleotide polymorphism"
FT replace(75962,T)
FT /tag= aj
FT /standard_name= "single nucleotide polymorphism"
FT replace(82853,A)
FT /tag= ak
FT /standard_name= "single nucleotide polymorphism"
FT replace(82930,C)
FT /tag= al
FT /standard_name= "single nucleotide polymorphism"
FT replace(88505,C)
FT /tag= am
FT /standard_name= "single nucleotide polymorphism"
FT replace(95970, .95971,A)
FT /tag= ao
FT /standard_name= "single nucleotide polymorphism"
FT replace(95970,A)
FT /tag= an
FT /standard_name= "single nucleotide polymorphism"
FT replace(96524,T)
FT /tag= ap
FT /standard_name= "single nucleotide polymorphism"
FT replace(100868,A)
FT /tag= aq
FT /standard_name= "single nucleotide polymorphism"
FT replace(102246,G)
FT /tag= ar
FT /standard_name= "single nucleotide polymorphism"
FT replace(107335,T)
FT /tag= as
FT /standard_name= "single nucleotide polymorphism"
FT replace(107921,T)
FT /tag= at
FT /standard_name= "single nucleotide polymorphism"
FT replace(110413,T)
FT /tag= au
FT /standard_name= "single nucleotide polymorphism"
FT replace(111600,G)
FT /tag= av
FT /standard_name= "single nucleotide polymorphism"
FT replace(114518,C)
FT /tag= aw
FT /standard_name= "single nucleotide polymorphism"
FT replace(114614,T)
FT /tag= ax
FT /standard_name= "single nucleotide polymorphism"
FT replace(124669,A)
FT /tag= ay
FT /standard_name= "single nucleotide polymorphism"

```
FT      /*tag= av
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(125409,A)
FT      /*tag= az
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(129447,A)
FT      /*tag= ba
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(135139,G)
FT      /*tag= bb
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(148111,T)
FT      /*tag= bc
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(200822,G)
FT      /*tag= bd
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(207967,G)
FT      /*tag= be
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(213624,C)
FT      /*tag= bf
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(215753,G)
FT      /*tag= bg
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(216081,A)
FT      /*tag= bh
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(218692,T)
FT      /*tag= bi
FT      /standard_name= "single nucleotide polymorphism"
FT      replace(218705,G)
FT      /*tag= bj
FT      /standard_name= "single nucleotide polymorphism"

Query Match      78.1%; Score 16.4; DB 8; Length 254366;
Best Local Similarity 94.4%; Pred. No. 7.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 AGATAGGCACCAATGA 21
        ||||| ||||| ||||| |||||
Db      228707 AGATATGCACACATGA 228690

RESULT 22
ADT05738
ID      ADT05738 standard; DNA; 308766 BP.
XX
AC      ADT05738;
XX
DT      02-DEC-2004 (first entry)
XX
DE      Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 774.
XX
KW      middle ear bacterial infection; nasopharynx bacterial infection; ds.
XX
OS      Haemophilus influenzae.
XX
PN      WO2004078949-A2.
XX
PD      16-SEP-2004.
XX
PF      05-MAR-2004; 2004WO-US007001.
XX
PR      06-MAR-2003; 2003US-0453134P.
XX
PA      (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI      Bakaletz LO, Munson RS, Dyer DW;
XX
DR      WPI; 2004-662422/64.
XX
```

```
PT      New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT      useful for treating or preventing NTHi bacterial infections of the middle
PT      ear and/or nasopharynx.
XX
PS      Disclosure; SEQ ID NO 774; 88pp; English.
XX
CC      The invention comprises nucleotide sequences (genes) from the genome of a
CC      nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC      sequences of the invention are useful for treating or preventing NTHi
CC      bacterial infections of the middle ear and/or nasopharynx. The present
CC      nucleic acid represents an NTHi DNA sequence that was used in the
CC      exemplification of the invention.
XX
SQ      Sequence 308766 BP; 97734 A; 61755 C; 55928 G; 93311 T; 0 U; 38 Other;

Query Match      78.1%; Score 16.4; DB 13; Length 308766;
Best Local Similarity 94.4%; Pred. No. 7.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 GAGATAGGCACCAATG 20
        ||||| ||||| ||||| |||||
Db      258587 GAGATAGGCACCAATG 258604

RESULT 23
AAI12660
ID      AAI12660 standard; DNA; 150 BP.
XX
AC      AAI12660;
XX
DT      30-MAR-1999 (first entry)
XX
DE      Human biallelic polymorphic DNA fragment stSG10266.
XX
KW      Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW      detection; phenotypic typing; characteristic; infection; hereditary;
KW      autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW      treatment; marker; ss.
XX
OS      Homo sapiens.
XX
PN      WO9820165-A2.
XX
PD      14-MAY-1998.
XX
PF      05-NOV-1997; 97WO-US020313.
XX
PR      06-NOV-1996; 96US-0030455P.
XX
PA      (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI      Lander ES, Wang D, Hudson T;
XX
DR      WPI; 1998-286974/25.
XX
PT      New isolated nucleic acid segments from the human genome - used for
PT      determining polymorphic forms for use in e.g. forensics, paternity
PT      testing or phenotypic typing for disease.
XX
PS      Claim 1; Page 278; 310pp; English.
XX
CC      AAX10269-X12937 are human DNA fragments which contain biallelic
CC      polymorphic markers which have been isolated using the primers
CC      represented in AAX09121-X10268. The base occupying the polymorphic site
CC      is indicated by the appropriate IUPAC-IUB ambiguity code. These fragments
CC      can be used in methods for determining polymorphic forms in an individual
CC      for use in e.g. forensics, paternity testing or for phenotypic typing for
CC      diseases such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan
CC      syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC      familial hypercholesterolemia, polycystic kidney disease, hereditary
CC      spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC      haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC      syndrome, osteogenesis imperfecta, acute intermittent porphyria,
```

CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases

XX
SQ Sequence 150 BP; 51 A; 23 C; 38 G; 37 T; 0 U; 1 Other;
Query Match 77.1%; Score 16.2; DB 2; Length 150;
Best Local Similarity 85.7%; Pred. No. 3.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACCAATGA 21
DB 73 GTGAGATGGCATGCAATGA 93
||||| ||||| ||||| |||||

RESULT 24
ABX87284/c
ID ABX87284 standard; cDNA; 294 BP.
XX
AC ABX87284;
XX
DT 24-APR-2003 (first entry)
XX
DE Corn ear-derived polynucleotide (cpd) #5744.
XX
KW Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW structural gene; functional gene; regulatory gene;
KW corn ear-specific profile; gene expression; gene expression;
KW hybrid plant; desirable trait expression; plant breeding program;
KW inheritance; desirable characteristic; growth; development;
KW disease resistance; environmental adaptability; quality; yield;
KW multigene trait; plant; gene; ss.
XX
OS Zea mays.
XX
PN US6476212-B1.
XX
PD 05-NOV-2002.
XX
PF 14-MAY-1999; 99US-00313294.
XX
PR 26-MAY-1998; 98US-0086722P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
DR WPI; 2003-208840/20.
XX
PT Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with growth,
PT development.
XX
PS Example; SEQ ID NO 5744; 390pp; English.
XX

CC The present invention relates to the isolation of corn ear-derived
CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC and SATMON023. Some of the cDps uniquely identify structural, functional,
CC and regulatory genes of corn ear. The polynucleotide sequences are
CC useful for detecting cDps in a sample, for producing a corn ear-specific
CC profile of gene transcription, for detecting altered gene expression in
CC inbred or hybrid plants, and for screening several molecules for specific
CC binding to the polynucleotide. The cDps are useful to identify, isolate,
CC or extend identical or related corn-ear nucleic acid sequences from DNA
CC libraries, and in nucleic acid amplification or hybridization techniques
CC to follow the expression of desirable traits through plant breeding
CC programs. Preferably, the cDps are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with growth

CC and development, disease resistance, environmental adaptability, quality,
CC and yield of corn. The cDps are also useful as molecular markers for
CC studying inheritance and multigene traits in a plant breeding program.
CC The cDps are useful for producing purified corn-ear polypeptides by
CC recombinant techniques. They are also useful in diagnostic assays to
CC detect or confirm conditions or diseases associated with abnormal levels
CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC polynucleotides (cdps) of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipdIDEntry.html
XX
SQ Sequence 294 BP; 75 A; 54 C; 58 G; 107 T; 0 U; 0 Other;
Query Match 77.1%; Score 16.2; DB 10; Length 294;
Best Local Similarity 85.7%; Pred. No. 3.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACCAATGA 21
DB 268 GTTAGATATGCACCAATTA 248
||||| ||||| ||||| |||||

RESULT 25
ACN52890
ID ACN52890 standard; cDNA; 602 BP.
XX
AC ACN52890;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton androecium tissue EST Clone ID: LIB3828-019-Q1-N6-A9, SEQ:7671.
XX
KW Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;
KW variety Nucotton33B; library LIB3828; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS Gossypium hirsutum.
XX
PN US2004123340-A1.
XX
PD 24-JUN-2004.
XX
PF 12-DEC-2001; 2001US-00021323.
XX
PR 14-DEC-2000; 2000US-0255619P.
XX
PA (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINC/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.
XX
PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;
XX
DR WPI; 2004-479808/45.
XX
PT New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.
XX
PS Claim 1; SEQ ID NO 7671; 34pp; English.
XX

CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN5220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nucotton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as

CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety Nucleon33B androecium tissue cDNA library (LIB3828). The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the US
 CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
 XX
 SQ Sequence 602 BP; 201 A; 124 C; 125 G; 152 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 13; Length 602;
 Best Local Similarity 85.7%; Pred. No. 3.9e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
 ||||| ||||| ||||| |||||
 Db 394 GTGACATAGCACATCAATGA 414

RESULT 26

ACN53918/c
 ID ACN53918 standard; cDNA; 624 BP.

XX ACN53918;

XX 02-DEC-2004 (first entry)

XX Cotton androecium tissue EST Clone ID: LIB3828-019-01-K6-A9, SEQ.8699.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;
 KW variety Nucleon33B; library LIB3828; molecular tag; molecular marker;
 KW genetic mapping; molecular mapping; seed germination; plant growth;
 KW plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

XX US2004123340-A1.

XX 24-JUN-2004.

XX 12-DEC-2001; 2001US-00021323.

XX 14-DEC-2000; 2000US-0255619P.

XX (DEIK/) DEIKMAN J.

PA (FENG/) FENG P C C.

PA (FINC/) FINCHER K L.

PA (ZIEG/) ZIEGLER T E.

XX Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its
 PT fragment, useful for isolating a variety of agronomically significant
 PT genes associated with plant growth, quality or yield, and as molecular
 PT tags to map genes.

XX Claim 1; SEQ ID NO 8699; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
 CC ACN5220-ACN63099). The ESTs were isolated from cDNA libraries generated
 CC from primed or non-primed seeds from variety DP50B, mature seeds from
 CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeceum
 CC tissue, developing fibres, carpel walls and septa from variety
 CC Nucleon33B. The invention also relates to substantially purified
 CC proteins or their fragments encoded by nucleic acid molecules of the
 CC invention, and to transformed plants having a nucleic acid construct
 CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
 CC molecular tags to isolate genetic regions, to isolate genes, to map
 CC genes, to determine gene function and to determine whether genes are
 CC members of a particular gene family. The nucleic acid molecules may be
 CC used for isolating a variety of agronomically significant genes
 CC associated with plant growth, quality, yield, and could also serve as
 CC links in metabolic and catabolic pathways. The nucleic acid molecules are
 CC also useful for identifying genes important in initiating and maintaining
 CC seed germination or that may be used to mitigate stresses encountered
 CC during seed germination. The ESTs additionally enable the acquisition of
 CC promoters and cis-regulatory elements which will be useful to express
 CC agronomically significant genes in these tissues and/or other tissues,
 CC and also permits the acquisition of molecular markers useful in breeding
 CC schemes, genetic and molecular mapping, and in cloning of agronomically
 CC significant genes. The nucleic acid molecules are further useful for
 CC detecting the expression level or pattern of a protein or mRNA and for
 CC detecting the presence or quantity of a protein by tissue printing. The
 CC present sequence represents a specifically claimed EST isolated from a
 CC cotton variety Nucleon33B androecium tissue cDNA library (LIB3828). The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the US
 CC patent office at seqdata.uspto.gov/sequence.html?DocID=US20040123340
 XX

SQ Sequence 624 BP; 177 A; 123 C; 140 G; 183 T; 0 U; 1 Other;

Query Match 77.1%; Score 16.2; DB 13; Length 624;

Best Local Similarity 85.7%; Pred. No. 3.9e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21

||||| ||||| ||||| |||||

Db 421 GTGACATAGCACATCAATGA 401

RESULT 27

ADSS4785/c

ID ADSS4785 standard; cDNA; 1964 BP.

XX AC ADSS4785;

XX 02-DEC-2004 (first entry)

XX Bacterial polynucleotide #6772.

XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polynucleotide; gene; ss.

OS Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

PA (HINK/) HINKLE G J.

PA (SLAT/) SLATER S C.
 PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 XX
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 30459; 122pp; English.
 XX
 CC The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polynucleotide used in
 CC the scope of the invention. Note: The sequence data for this patent did
 CC not form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 SQ Sequence 1964 BP; 569 A; 375 C; 341 G; 679 T; 0 U; 0 Other;
 Query Match 77.1%; Score 16.2; DB 13; Length 1964;
 Best Local Similarity 85.7%; Pred. No. 4.6e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GTGAGATAGGCACACAACTGA 21
 |||||
 Db 1928 GTGACATAGGCATAACTATGA 1908
 RESULT 28
 ADQ63615
 ID ADQ63615 standard; cDNA; 2210 BP.
 XX
 AC ADQ63615;
 XX
 XX 07-OCT-2004 (first entry)
 XX
 DE Novel human cDNA sequence #776.
 XX
 KW ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
 KW cyostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
 KW neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
 KW cancer.
 XX
 OS Homo sapiens.
 XX
 PN EPI440981-A2.
 XX
 XX 28-JUL-2004.
 XX
 XX 21-JAN-2004; 2004EP-00001196.

PR 21-JAN-2003; 2003JP-00102206.
 PR 09-MAY-2003; 2003JP-00131392.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Nagai K, Irie R;
 XX WPI; 2004-535376/52.
 DR P-PSDB; ADQ65803.
 XX
 XX Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX
 XX Claim 1; SEQ ID NO 776; 2449pp; English.
 XX
 CC The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.
 XX
 SQ Sequence 2210 BP; 489 A; 563 C; 528 G; 630 T; 0 U; 0 Other;
 Query Match 77.1%; Score 16.2; DB 12; Length 2210;
 Best Local Similarity 85.7%; Pred. No. 4.7e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GTGAGATAGGCACACAACTGA 21
 |||||
 Db 326 GTGAGATAGGCACATCATGTA 346
 RESULT 29
 ABL23258
 ID ABL23258 standard; DNA; 4045 BP.
 XX
 AC ABL23258;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21247.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 XX Claim 1; SEQ ID NO 21247; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 4045 BP; 1030 A; 963 C; 961 G; 1091 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 4; Length 4045;

Best Local Similarity 85.7%; Pred. No. 5.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21

||||| ||||| ||||| |||||

Db 998 GTGAGAAAGGCACGAAATGA 1018

RESULT 30

ADD48369

ID ADD48369 standard; DNA; 4576 BP.

XX ADD48369;

XX AC

XX 02-DEC-2004 (revised)

DT 29-JAN-2004 (first entry)

XX

DE Human gene NM_000222, SEQ ID NO 14070.

XX

KW Human; ds; gene; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

OS Unidentified.

XX

PN WO2003016475-A2.

XX

XX 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR

PR 01-NOV-2001; 2001US-0346382P.

PR

PR 26-NOV-2001; 2001US-0333347P.

XX

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

XX Woolf C, D'urso D, Befort K, Costigan M;

XX

XX WPI; 2003-268312/26.

DR GENBANK; NM_000222.

XX

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX

PS Example 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human DNA (described in Table 3 of
CC the specification) which encodes one of the polypeptides of the invention
CC which is differentially expressed during pain. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 4576 BP; 1289 A; 960 C; 1003 G; 1324 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 10; Length 4576;

Best Local Similarity 85.7%; Pred. No. 5.2e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21

||||| ||||| ||||| |||||

Db 2021 GTGAGATAGGCTCAACATAGA 2041

RESULT 31

ADE55806

ID ADE55806 standard; DNA; 4576 BP.

XX

AC ADE55806;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human gene NM_000222, SEQ ID NO 1637.

XX

KW Human; ds; gene; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

PN WO2003016475-A2.

XX

XX 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

PR 14-AUG-2001; 2001US-0312147P.

PR

PR 01-NOV-2001; 2001US-0346382P.

PR

PR 26-NOV-2001; 2001US-0333347P.

XX

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

XX Woolf C, D'urso D, Befort K, Costigan M;

XX

XX WPI; 2003-268312/26.

DR GENBANK; NM_000222.

XX

PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX

PS Claim 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human DNA (shown in Table 2 of the
 CC specification) which encodes one of the polypeptides of the invention
 CC which is differentially expressed during pain. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 4576 BP; 1289 A; 960 C; 1003 G; 1324 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 10; Length 4576;
 Best Local Similarity 85.7%; Pred. No. 5.2e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
 |||||
 DB 2021 GTGAGATAGGCACACAAATGA 2041

RESULT 32

AAA75080/c
 ID AAA75080 standard; DNA; 44848 BP.

XX AC AAA75080;

XX DT 15-JAN-2001 (first entry)

XX DE Nucleotide sequence of the human heparanase gene.

XX KW Human; heparanase; gene therapy; tumour; inflammation; autoimmunity;
 KW heparin-binding growth factor; cytokine; neurodegenerative plaque;
 KW wound healing; infection; burn; angiogenesis; restenosis;
 KW atherosclerosis; inflammation; neurodegenerative disease;
 KW Gerstmann-Straussler Syndrome; Creutzfeldt-Jakob disease; ds.

XX OS Homo sapiens.

XX PN WO200052178-A1.

XX PD 08-SEP-2000.

XX PF 14-FEB-2000; 2000WO-US003542.

XX PR 01-MAR-1999; 99US-00258892.

XX PA (INSI-) INSIGHT STRATEGY & MARKETING LTD.
 PA (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 PA (FRIE/) FRIEDMAN M.M.

XX PI Pecker I, Vlodavsky I, Feinstein E;

XX DR WPI; 2000-579289/54.

XX PT New polynucleotides encoding a polypeptide having heparanase activity,
 PT useful in wound healing and in gene therapy, particularly in treating
 PT tumor, inflammation, autoimmunity, neurodegenerative diseases.

XX PS

XX Claim 9; Page 131-143; 152pp; English.

XX CC The present sequence represents a human gene which encodes a protein with
 CC heparanase catalytic activity. The heparanase (hpa) polynucleotide is
 CC useful in gene therapy, particularly in treating tumour, inflammation or
 CC autoimmunity. Particularly, the polynucleotide is useful in modulating
 CC the bioavailability of heparin-binding growth factors, cellular responses
 CC to heparin-binding growth factors (e.g. bFGF) and cytokines (e.g.
 CC interleukin (IL)-8), cell interaction with plasma lipoproteins, cellular
 CC susceptibility to certain viral and some bacterial and protozoa
 CC infections, or disintegration of neurodegenerative plaques. The
 CC polynucleotide is also useful in wound healing (e.g. thermal, chemical or
 CC radiation burns), and in the treatment of angiogenesis, restenosis, or
 CC atherosclerosis, inflammation, neurodegenerative diseases (Gerstmann-
 CC Straussler Syndrome or Creutzfeldt-Jakob disease), and some viral,
 CC bacterial or protozoa infections

XX SQ Sequence 44848 BP; 12560 A; 9646 C; 8930 G; 13712 T; 0 U; 0 Other;

Query Match 77.1%; Score 16.2; DB 3; Length 44848;

Best Local Similarity 85.7%; Pred. No. 7.2e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21

DB 24187 GTTAGATTGGCACCACAAATGA 24167

RESULT 33

ADG88832/c

ID ADG88832 standard; DNA; 44848 BP.

XX AC ADG88832;

XX DT 11-MAR-2004 (first entry)

XX DE Human hpa genomic DNA.

XX KW Wound healing; heparanase; ulcer; burn; laceration; surgical incision;
 KW necrosis; pressure wound; diabetic ulcer; angiogenesis; human; therapy;
 KW gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 2743..41863

XX FT /*tag= a

XX FT /product= "Human hpa protein"

XX PN US2003161823-A1.

XX PD 28-AUG-2003.

XX PF 14-JAN-2003; 2003US-00341582.

XX PR 31-AUG-1998; 98WO-US017954.

XX PR 01-MAR-1999; 99US-00258892.

XX PR 06-FEB-2001; 2001US-00776874.

XX PR 05-SEP-2001; 2001WO-IL000830.

XX PR 19-NOV-2001; 2001US-00988113.

XX PA (ILAN/) ILAN N.

XX PA (VLOD/) VLODAVSKY I.

XX PA (YACO/) YACOBY-ZEEVI O.

XX PA (PECK/) PECKER I.

XX PA (FEIN/) FEINSTEIN E.

XX PI Ilan N, Vlodavsky I, Yacoby-Zeevi O, Pecker I, Feinstein E;

XX DR WPI; 2003-897910/82.

XX DR P-PSDB; ADG88800.

PT Composition for treating a wound comprising recombinant heparanase is
PT useful to induce or accelerate wound healing and induce or accelerate
XX angiogenesis.
PS Claim 4; SEQ ID NO 42; 143pp; English.
XX
CC The present invention relates to methods and compositions for inducing
CC and/or accelerating wound healing via the catalytic activity of
CC heparanase. The invention is used to induce or accelerate a healing
CC process, particularly of an ulcer, burn, laceration, surgical incision,
CC necrosis, pressure wound, diabetic ulcer and to induce or accelerate
CC angiogenesis. The present sequence is human hpa genomic DNA.
XX
SQ Sequence 44848 BP; 12560 A; 9646 C; 8930 G; 13712 T; 0 U; 0 Other;
Query Match 77.1%; Score 16.2; DB 10; Length 44848;
Best Local Similarity 85.7%; Pred. No. 7.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GTGAGATAGGCACCAACATGA 21
DB 24187 GTTAGATTGGCACCACATGA 24167
RESULT 34
ADL16411/c
ID ADL16411 standard; DNA; 44848 BP.
XX
AC ADL16411;
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Human heparanase genomic DNA.
DE
XX
XX Human; ds; heparanase; gene; heparanase-dependent cancer; cancer;
KW autoimmune reaction; inflammation; chromosome 4.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH promoter 2635..2635
FT /tag= a
FT /note= "Minimal promoter region. Claimed in claim 15"
FT 5'UTR 2635..2742
FT /tag= b
FT /note= "Claimed in claim 19"
FT 3'UTR 41864..41890
FT /tag= c
FT /note= "Claimed in claim 20"
FT
XX
XX US2003236215-A1.
XX
XX 25-DEC-2003.
XX
XX 09-JUN-2003; 2003US-00456573.
XX
XX 31-AUG-1998; 98WO-US017954.
XX 01-MAR-1999; 99US-00258892.
XX 08-NOV-1999; 99US-00435739.
XX
XX (INSI-) INSIGHT STRATEGY & MARKETING LTD.
XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
XX Pecker I, Vlodavsky I, Feinstein E;
XX
XX WPI; 2004-070610/07.
XX
XX New antisense oligonucleotide hybridizable with a polynucleotide encoding
XX a polypeptide with heparanase activity, useful for treating diseases such
XX as cancer and autoimmune disorders.
XX
XX Claim 2; SEQ ID NO 42; 108pp; English.
XX

CC The invention relates to an antisense oligonucleotide (ASO) comprising a
CC polynucleotide or a polynucleotide analogue of at least 10 bases being
CC hybridisable in vivo, under physiological conditions, with a portion of
CC a polynucleotide strand encoding a polypeptide having heparanase
CC catalytic activity. Also included are a method of in vivo downregulating
CC heparanase activity (comprising administering the ASO in vivo), a method
CC of treating a subject suffering from a pathological condition
CC (characterised by heparanase activity, comprising administering ASO to
CC the subject), a pharmaceutical composition comprising the ASO and a
CC carrier, an antisense nucleic acid construct (comprising a promoter
CC sequence and a polynucleotide sequence directing the synthesis of an
CC antisense RNA sequence of at least 10 bases being hybridisable in vivo,
CC under physiological conditions, with a polynucleotide strand encoding a
CC polypeptide having heparanase catalytic activity), a method of in vivo
CC downregulating heparanase activity (comprising administering in vivo the
CC antisense nucleic acid construct), a pharmaceutical composition
CC comprising the antisense nucleic acid construct and a carrier, and an
CC antisense oligonucleotide comprising a polynucleotide or a polynucleotide
CC analogue of at least 10 bases being hybridisable in vivo, under
CC physiological conditions, with a portion of a polynucleotide strand being
CC characterised by forming at least a portion of an untranslated region
CC (UTR) for a polynucleotide strand encoding a polypeptide having
CC heparanase catalytic activity. The methods and compositions of the
CC present invention are useful for the prevention and/or treatment of
CC diseases or conditions associated with aberrant heparanase activity, such
CC as heparanase-dependent cancer, cancer, autoimmune reaction and
CC inflammation. The gene for human heparanase is located on chromosome 4.
CC The present sequence is the human heparanase gene.
XX
XX Sequence 44848 BP; 12560 A; 9646 C; 8930 G; 13712 T; 0 U; 0 Other;
SQ

Query Match 77.1%; Score 16.2; DB 12; Length 44848;
Best Local Similarity 85.7%; Pred. No. 7.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GTGAGATAGGCACCAACATGA 21
DB 24187 GTTAGATTGGCACCACATGA 24167
RESULT 35
ADM48748/c
ID ADM48748 standard; DNA; 44848 BP.
XX
AC ADM48748;
XX
XX 03-JUN-2004 (first entry)
DT
XX
XX Human hpa genomic DNA.
XX
XX Transgenic animal; heparanase; cancer; viral infection; restenosis;
KW neurodegenerative disease; atherosclerosis; pulmonary disorder; hpa;
KW human; gene; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH exon 2635..2969
FT /tag= a
FT /number= 1
FT CDS 2743..41863
FT /tag= b
FT /product= "Hpa protein"
FT intron 2970..15141
FT /tag= c
FT /number= 1
FT exon 15142..15287
FT /tag= d
FT /number= 2
FT intron 15288..18035
FT /tag= e
FT /number= 2
FT exon 18036..18161
FT

FT FT /*tag= f
 FT /number= 3
 FT 18162..24214
 FT /*tag= g
 FT /number= 3
 FT 24215..24388
 FT /*tag= h
 FT /number= 4
 FT 24389..26600
 FT /*tag= i
 FT /number= 4
 FT 26601..26769
 FT /*tag= j
 FT /number= 5
 FT 26770..27413
 FT /*tag= k
 FT /number= 5
 FT 27414..27461
 FT /*tag= l
 FT /number= 6
 FT 27462..27996
 FT /*tag= m
 FT /number= 6
 FT 27997..28090
 FT /*tag= n
 FT /number= 7
 FT 28091..28540
 FT /*tag= o
 FT /number= 7
 FT 28541..28647
 FT /*tag= p
 FT /number= 8
 FT 28648..31173
 FT /*tag= q
 FT /number= 8
 FT 31174..31288
 FT /*tag= r
 FT /number= 9
 FT 31289..35227
 FT /*tag= s
 FT /number= 9
 FT 35228..35346
 FT /*tag= t
 FT /number= 10
 FT 35347..36389
 FT /*tag= u
 FT /number= 10
 FT 36390..36536
 FT /*tag= v
 FT /number= 11
 FT 36537..41703
 FT /*tag= w
 FT /number= 11
 FT 41704..41890
 FT /*tag= x
 FT /number= 12
 FT FT
 XX US2003217375-A1.
 XX
 XX 20-NOV-2003.
 XX
 XX 24-FEB-2003; 2003US-00371218.
 XX
 XX 31-AUG-1998; 98WO-US017954.
 XX 01-MAR-1999; 99US-00258892.
 XX 06-FEB-2001; 2001US-00776874.
 XX 19-NOV-2001; 2001US-00988113.
 XX
 PA (ZCHA/) ZCHARIA E.
 PA (VLOD/) VLODAVSKY I.
 PA (METZ/) METZGER S.
 PA (PECK/) PECKER I.
 PA (ILAN/) ILAN N.

PA (CHAJ/) CHAJEK-SHAUL T.
 PA (GOLD/) GOLDSCHMIDT O.
 XX
 PI Zcharia E, Vlodavsky I, Metzger S, Pecker I, Ilan N;
 PI Chajek-Shaul T, Goldshmidt O;
 XX
 XX WPI: 2004-021918/02.
 DR P-PSDB; ADM48759.
 XX
 FT New transgenic non-human animal expressing heparinase, useful as models
 FT for human disease, such as cancers, viral infection, neurodegenerative
 FT diseases, restenosis, atherosclerosis and pulmonary disorders.
 XX
 PS Example 10; SEQ ID NO 42; 106pp; English.
 XX
 CC The present invention relates to a transgenic non-human animal whose
 CC genome comprises an exogenous polynucleotide sequence, including a
 CC promoter active in tissues of the non-human, a region encoding a human
 CC heparanase, where the promoter and the region encoding human heparanase
 CC are operably linked in the exogenous polynucleotide such that human
 CC heparanase is expressed in at least a portion of the cells of the non-
 CC human animal. The methods and compositions of the present invention are
 CC useful for the production of transgenic animals expressing heparanase, to
 CC be used as models for human diseases such as cancers, viral infection,
 CC restenosis, neurodegenerative diseases, atherosclerosis and pulmonary
 CC disorders. The present sequence is human hpa genomic DNA used in the
 CC exemplification of the invention.
 XX
 SQ Sequence 44848 BP; 12560 A; 9646 C; 8930 G; 13712 T; 0 U; 0 Other;
 Query Match 77.1%; Score 16.2; DB 12; Length 44848;
 Best Local Similarity 85.7%; Pred. No. 7.2e+02;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GTGACATAGGCACACATGA 21
 |||||
 DB 24187 GTTAGATTGCACCATGA 24167
 RESULT 36
 ACA60949/c
 ID ACA60949 standard; DNA; 55827 BP.
 XX
 AC ACA60949;
 XX
 DT 11-AUG-2003 (first entry)
 XX
 DE DNA encoding human carboxypeptidase.
 XX
 KW Human; gene; carboxypeptidase; inflammation; cancer; arteriosclerosis;
 KW neurodegenerative disease; protease; ds; single nucleotide polymorphism;
 KW SNP.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT variation replace(858,T)
 FT /*tag= u
 FT /note= "Single nucleotide polymorphism"
 FT 2017..53409
 FT CDS
 FT /*tag= a
 FT /product= "Carboxypeptidase"
 FT 2017..2090
 FT exon
 FT /*tag= b
 FT /number= 1
 FT 2091..19206
 FT intron
 FT /*tag= c
 FT /number= 1
 FT variation replace(2122,T)
 FT /*tag= v
 FT /note= "Single nucleotide polymorphism"
 FT variation replace(4088,A)
 FT /*tag= w

FT	variation	/note= "Single nucleotide polymorphism"	FT	intron	28232.. .33052
FT		/tag= x	FT		/tag= k
FT		/number= 5	FT	variation	replace(28476,C)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= ao
FT		replace(4416.. .4418,TA)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= y	FT		replace(29404,G)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= ap
FT		replace(4434,G)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= z	FT		replace(31047,T)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= aq
FT		replace(4969,C)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= aa	FT		replace(31445,C)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= ar
FT		replace(5868,A)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= ab	FT		replace(31447,C)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= ag
FT		replace(6254,C)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= ac	FT		replace(31600,C)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= at
FT		replace(10171,C)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= ad	FT		replace(31714,C)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= au
FT		replace(10452,A)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= ae	FT		replace(31715,T)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= av
FT		replace(11613,A)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= af	FT		replace(32193,G)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= aw
FT		replace(12130,T)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= ag	FT		replace(32341,A)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= ay
FT		replace(17867.. .17869,GA)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= ah	FT		replace(32341,G)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= ax
FT		replace(18243,G)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= ai	FT		replace(32561,G)
FT	exon	/note= "Single nucleotide polymorphism"	FT		/tag= az
FT		19207.. .19282	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= d	FT		replace(32600.. .32602,AA)
FT	intron	/number= 2	FT		/tag= ba
FT		19283.. .22683..	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= e	FT		replace(32642,T)
FT		/number= 2	FT	variation	/tag= bb
FT	variation	replace(22450,C)	FT		/note= "Single nucleotide polymorphism"
FT		/tag= aj	FT	variation	replace(32793.. .32795,TG)
FT		/note= "Single nucleotide polymorphism"	FT		/tag= bc
FT	exon	22684.. .22808	FT		/note= "Single nucleotide polymorphism"
FT		/tag= f	FT	exon	33053.. .33157
FT		/number= 3	FT		/tag= l
FT	intron	22809.. .24477	FT		/number= 6
FT		/tag= g	FT	variation	replace(33071,A)
FT		/number= 3	FT		/tag= bd
FT	variation	replace(23003,T)	FT	intron	/note= "Single nucleotide polymorphism"
FT		/tag= ak	FT		33158.. .42288
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= m
FT		replace(24055.. .24057,GA)	FT	variation	/number= 6
FT		/tag= al	FT		replace(34721,T)
FT	variation	/note= "Single nucleotide polymorphism"	FT		/tag= be
FT		replace(24132,T)	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= am	FT		replace(35304,A)
FT	exon	/note= "Single nucleotide polymorphism"	FT		/tag= bf
FT		24478.. .24586	FT	variation	/note= "Single nucleotide polymorphism"
FT		/tag= h	FT		replace(35425,C)
FT		/number= 4	FT	variation	/tag= bg
FT	intron	24587.. .28129	FT		/note= "Single nucleotide polymorphism"
FT		/tag= i	FT		replace(36050,G)
FT		/number= 4	FT	variation	/tag= bh
FT	variation	replace(25713,A)	FT		/note= "Single nucleotide polymorphism"
FT		/tag= an	FT	variation	replace(36291,G)
FT		/note= "Single nucleotide polymorphism"	FT		/tag= bi
FT	exon	28130.. .28231	FT		/note= "Single nucleotide polymorphism"
FT		/tag= j	FT	exon	42289.. .42382
FT		/number= 5	FT		

```
FT FT      /*tag= n
FT FT      /number= 7
FT FT      42383..48643
FT FT      /*tag= o
FT FT      /number= 7
FT FT      48644..48846
FT FT      /*tag= p
FT FT      /number= 8
FT FT      48847..51176
FT FT      /*tag= q
FT FT      /number= 8
FT FT      51177..51264
FT FT      /*tag= r
FT FT      /number= 9
FT FT      51265..53224
FT FT      /*tag= s
FT FT      /number= 9
FT FT      53225..53406
FT FT      /*tag= t
FT FT      /number= 10
XX
PN US2003017574-A1.

Query Match      77.1%; Score 16.2; DB 8; Length 55827;
Best Local Similarity 85.7%; Pred. No. 7.4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  GTGAGTAGGCACACAAATGA 21
        ||||| ||||| |||||
DB      27659 GTGAGATTGGCACAAATGATGA 27639

RESULT 37
ABX13671/C
ID ABX13671 standard; DNA; 55827 BP.
XX
AC ABX13671;
XX
DT 14-FEB-2003 (first entry)
XX
DE Human protease gene.
XX
KW Human; gene; ds; protease; proteolytic degradation; proteolysis;
KW proliferation; differentiation; signalling; therapeutic; gene therapy;
KW protein therapy; diagnostic; immune response; vaccine; inflammation;
KW cancer; arteriosclerosis; degenerative disorder; chromosome 13;
KW single nucleotide polymorphism; SNP.
XX
OS Homo sapiens.
XX
FH Key      Location/Qualifiers
FH variation replace(858,T)
FT      /*tag= u
FT      /standard_name= "Single nucleotide polymorphism"
FT      2017..53409
FT      /*tag= a
FT      /product= "Protease"
FT      2017..2090
FT      /*tag= b
FT      /number= 1
FT      2091..19206
FT      /*tag= c
FT      /number= 1
FT      /cons_splice= (5'site:yes,3'site:no)
FT      replace(2122,T)
FT      /*tag= v
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4088,A)
FT      /*tag= w
FT      /standard_name= "Single nucleotide polymorphism"
FT      replace(4240,A)
FT      /*tag= x
FT      /standard_name= "Single nucleotide polymorphism"
FT      28232..33052

FT FT      replace(4416..4418,TA)
FT FT      /*tag= y
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4434,G)
FT FT      /*tag= z
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(4569,C)
FT FT      /*tag= aa
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(5868,C)
FT FT      /*tag= ab
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(6254,C)
FT FT      /*tag= ac
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(10171,C)
FT FT      /*tag= ad
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(10452,A)
FT FT      /*tag= ae
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(11613,A)
FT FT      /*tag= af
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(12130,T)
FT FT      /*tag= ag
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(17867..17868,GAA)
FT FT      /*tag= ah
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(18243,G)
FT FT      /*tag= ai
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      19207..19282
FT FT      /*tag= d
FT FT      /number= 2
FT FT      19283..22683
FT FT      /*tag= e
FT FT      /number= 2
FT FT      /cons_splice= (5'site:yes,3'site:no)
FT FT      replace(22450,C)
FT FT      /*tag= aj
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      22684..22808
FT FT      /*tag= f
FT FT      /number= 3
FT FT      22809..24477
FT FT      /*tag= g
FT FT      /number= 3
FT FT      /cons_splice= (5'site:yes,3'site:no)
FT FT      replace(23003,T)
FT FT      /*tag= ak
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(24055..24056,GAA)
FT FT      /*tag= al
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      replace(24132,C)
FT FT      /*tag= am
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      24478..24586
FT FT      /*tag= h
FT FT      /number= 4
FT FT      24587..28129
FT FT      /*tag= i
FT FT      /number= 4
FT FT      /cons_splice= (5'site:no,3'site:no)
FT FT      replace(25713,A)
FT FT      /*tag= an
FT FT      /standard_name= "Single nucleotide polymorphism"
FT FT      28130..28231
FT FT      /*tag= j
FT FT      /number= 5
FT FT      28232..33052
```


CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The
CC present nucleic acid represents a human autoimmune disease-related
CC genomic DNA sequence of the invention. NOTE: The present sequence is not
CC shown in the specification, but has been retrieved from the WIPO website.
XX
SQ Sequence 58337 BP; 18163 A; 11254 C; 11433 G; 17380 T; 0 U; 107 Other;
Query Match 77.1%; Score 16.2; DB 13; Length 58337;
Best Local Similarity 85.7%; Pred. No. 7.5e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GTGAGATAGGCACACAAATGA 21
Db 32807 GTGAGATTGGCACAAATGATGA 32827
RESULT 39
AD536462/c
ID ADS36462 standard; DNA; 64423 BP.
XX
AC ADS36462;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human autoimmune disease-related genomic DNA sequence - SEQ ID 1676.
XX
KW single nucleotide polymorphism detection; SNP detection;
KW rheumatoid arthritis; type 1 diabetes; multiple sclerosis;
KW systemic lupus erythematosus; inflammatory bowel disease; psoriasis;
KW thyroiditis; celiac disease; pernicious anaemia; asthma; vitiligo;
KW glomerulonephritis; Grave's disease; myocarditis; Sjogren's disease;
KW primary systemic vasculitis; ds.
XX
OS Homo sapiens.
XX
PN WO2004083403-A2.
XX
PD 30-SEP-2004.
XX
PF 18-MAR-2004; 2004WO-US0008461.
XX
PR 18-MAR-2003; 2003US-0455444P.
XX
PR 25-APR-2003; 2003US-0465241P.
XX
PA (APPL-) APPLERA CORP.
XX
PI Cargill M, Begovich AB, Alexander HC;
XX
XX WPI; 2004-728480/71.
XX
XX New isolated nucleic acid molecule comprises at least 8 contiguous
XX nucleotides where one of the nucleotides is a single nucleotide
XX polymorphism (SNP), useful for diagnosing or treating autoimmune
XX diseases, e.g. rheumatoid arthritis.
XX
PS Claim 16; SEQ ID NO 1676; 123pp; English.
XX
CC The invention comprises amino acid and coding sequences containing
CC genetic polymorphisms associated with an altered risk of developing an
CC autoimmune disease (e.g. rheumatoid arthritis). The invention further
CC comprises a method of identifying an individual that has an altered risk
CC of developing an autoimmune disease, comprising detecting a single
CC nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA
CC and protein sequences of the invention are useful for diagnosing and
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious
CC anaemia, asthma, vitiligo, glomerulonephritis, Grave's disease,
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The

CC present nucleic acid represents a human autoimmune disease-related
CC genomic DNA sequence of the invention. NOTE: The present sequence is not
CC shown in the specification, but has been retrieved from the WIPO website.
XX
SQ Sequence 64423 BP; 19548 A; 12552 C; 12463 G; 19737 T; 0 U; 123 Other;
Query Match 77.1%; Score 16.2; DB 13; Length 64423;
Best Local Similarity 85.7%; Pred. No. 7.6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GTGAGATAGGCACACAAATGA 21
Db 31708 GTGAGATTGGCACAAATGATGA 31688
RESULT 40
AAL61326
ID AAL61326 standard; DNA; 91000 BP.
XX
XX AAL61326;
XX
DT 22-SEP-2003 (first entry)
XX
DE Human farnesoid X receptor (FXR) DNA #2.
XX
KW Human; farnesoid X receptor; FXR; cardiovascular disease; gene therapy;
KW atherosclerosis; hypercholesterolaemia; NR1H4; bile acid receptor; BAR;
KW retinoid X receptor-interacting protein 14; RXR14; gene; ds.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Intron 576..19018
FT /tag= a
FT /number= 1
FT exon 19019..19153
FT /tag= b
FT /number= 2
FT Intron 19154..19785
FT /tag= c
FT /number= 2
FT exon 19786..19918
FT /tag= d
FT /number= 3
FT Intron 19919..37293
FT /tag= e
FT /number= 3
FT exon 37294..37659
FT /tag= f
FT /number= 4
FT Intron 37660..58973
FT /tag= g
FT /number= 4
FT exon 58974..59114
FT /tag= h
FT /number= 5
FT Intron 59115..61405
FT /tag= i
FT /number= 5
FT exon 61406..61539
FT /tag= j
FT /number= 6
FT Intron 61540..63027
FT /tag= k
FT /number= 6
FT exon 63028..63126
FT /tag= l
FT /number= 7
FT Intron 63127..63463
FT /tag= m
FT /number= 7
FT exon 63464..63563
FT /tag= n

FT intron /number= 8
FT 63564. .67187
FT /*tag= 0
FT /number= 8
FT 67188. .67334
FT /*tag= P
FT /number= 9
FT 67335. .87922
FT /*tag= q
FT /number= 9
FT 87923. .88036
FT /*tag= r
FT /number= 10
FT 88037. .89288
FT /*tag= s
FT /number= 10
XX
XX WO2003044167-A2.
XX
XX 30-MAY-2003.
XX
XX 13-NOV-2002; 2002WO-US036691.
XX
XX 15-NOV-2001; 2001US-00002491.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Monia BP, Watt AT;
XX
XX WPI; 2003-468767/44.
XX
XX New antisense oligonucleotides for modulating human farnesoid X receptor
XX (FXR) expression, useful for treating conditions associated with FXR in
XX humans, e.g. cardiovascular disease, atherosclerosis or
XX hypercholesterolemia.
XX
XX Example 15; Page 84-110; 127pp; English.
XX
XX The invention relates to antisense compounds, compositions and methods
XX for modulating the expression of human farnesoid X receptor (FXR). FXR is
XX also known as NR1H4, retinoid X receptor-interacting protein 14 (RIP14)
XX and bile acid receptor (BAR). The antisense oligonucleotide is useful for
XX inhibiting the expression of human FXR in cells or tissues. It is
XX particularly useful for treating or preventing a disease or condition
XX associated with FXR in a human, e.g. cardiovascular disease,
XX atherosclerosis or hypercholesterolemia. The antisense compound is
XX useful for diagnostics, therapeutics, prophylaxis, or as research
XX reagents or kits. It is also used in gene therapy. The present sequence
XX is human FXR DNA
XX
SQ Sequence 91000 BP; 27125 A; 17597 C; 18434 G; 27544 T; 0 U; 300 Other;,
Query Match 77.1%; Score 16.2; DB 9; Length 91000;
Best Local Similarity 85.7%; Pred. No. 7.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GTGAGATAGGCACACAAATCA 21
||| ||||| ||||| ||||| |||||
Db 48586 GTAAGATAGGCACACAAATGA 48606
Search completed: June 4, 2005, 07:29:09
Job time : 163.178 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:48:59 ; Search time 46.7219 Seconds
(without alignments)
735.454 Million cell updates/sec

Title: US-09-674-277-27

Perfect score: 21

Sequence: 1 GTGAGATAGGCACCAATGA 21

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	17.8	84.8	239527	4	US-09-949-016-15980
C 2	16.8	80.0	294836	4	US-09-949-016-15974
C 3	16.4	78.1	601	4	US-09-949-016-40710
C 4	16.4	78.1	254366	4	US-09-822-871-3
C 5	16.4	78.1	462589	4	US-09-949-016-12900
C 6	16.4	78.1	476044	4	US-09-949-016-12412
C 7	16.4	78.1	1830121	4	US-09-557-884-1
C 8	16.4	78.1	1830121	4	US-09-643-990A-1
C 9	16.2	77.1	234	4	US-09-313-294A-5744
C 10	16.2	77.1	601	4	US-09-949-016-78348
C 11	16.2	77.1	601	4	US-09-949-016-167629
C 12	16.2	77.1	601	4	US-09-949-016-167736
C 13	16.2	77.1	699	4	US-09-248-796A-11382
C 14	16.2	77.1	44848	4	US-09-435-739-42
C 15	16.2	77.1	44848	4	US-09-988-113-42
C 16	16.2	77.1	55827	3	US-09-813-133A-3
C 17	16.2	77.1	55827	3	US-10-212-877-3
C 18	16.2	77.1	74096	4	US-09-949-016-11785
C 19	16.2	77.1	74097	4	US-09-949-016-16239
C 20	16.2	77.1	78157	4	US-09-949-016-16466
C 21	16.2	77.1	78157	4	US-09-949-016-16467
C 22	16.2	77.1	94019	4	US-09-949-016-16467
C 23	16.2	77.1	187169	4	US-09-949-016-12776
C 24	16.2	77.1	191569	4	US-09-949-016-15940
C 25	16.2	77.1	276237	4	US-09-949-016-17504
C 26	16.2	77.1	784019	4	US-09-949-016-14033
C 27	16.2	77.1	828152	4	US-09-949-016-12777

28	16	76.2	601	4	US-09-949-016-76772	Sequence 76772, A
C 29	16	76.2	57638	4	US-09-949-016-17000	Sequence 17000, A
C 30	16	76.2	301828	4	US-09-949-016-13969	Sequence 13969, A
C 31	15.8	75.2	1497	4	US-09-107-532A-3337	Sequence 3337, Ap
C 32	15.8	75.2	12995	4	US-09-949-016-14249	Sequence 14249, A
C 33	15.8	75.2	12995	4	US-09-949-016-14250	Sequence 14250, A
C 34	15.8	75.2	12995	4	US-09-949-016-14251	Sequence 14251, A
C 35	15.8	75.2	12995	4	US-09-949-016-14252	Sequence 14252, A
C 36	15.8	75.2	24496	4	US-09-949-016-11823	Sequence 11823, A
C 37	15.8	75.2	24497	4	US-09-949-016-14253	Sequence 14253, A
C 38	15.8	75.2	48974	3	US-08-920-422-17	Sequence 17, Appl
C 39	15.8	75.2	422592	4	US-09-949-016-14182	Sequence 14182, A
C 40	15.8	75.2	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 41	15.8	75.2	1664976	4	US-09-692-570-1	Sequence 1, Appli
C 42	15.4	73.3	838	4	US-09-774-639-21	Sequence 21, Appl
C 43	15.4	73.3	88245	4	US-09-949-016-13835	Sequence 13835, A
C 44	15.2	72.4	156	4	US-09-493-795B-261	Sequence 261, App
C 45	15.2	72.4	196	3	US-08-651-155B-181	Sequence 181, App

ALIGNMENTS

RESULT 1

US-09-949-016-15980/c
; Sequence 15980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15980
; LENGTH: 239527
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(239527)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15980

Query Match 84.8%; Score 17.8; DB 4; Length 239527;
Best Local Similarity 90.5%; Pred. No. 40;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GTGAGATAGGCACCAATGA 21
Db 135081 GTGAGATAGGCACCAATGA 135061

RESULT 2

US-09-949-016-15974/c
; Sequence 15974, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15974
; LENGTH: 294836
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(294836)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15974
```

```
Query Match      80.0%; Score 16.8; DB 4; Length 294836;
Best Local Similarity 90.0%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY      1  GTGAGATAGGCACAAATG 20
          ||||| ||||| ||||| |||||
Db      167532 GTCAATGGGCACAAATG 167513
```

RESULT 3

```
US-09-949-016-40710
; Sequence 40710, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40710
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-40710
```

```
Query Match      78.1%; Score 16.4; DB 4; Length 601;
Best Local Similarity 94.4%; Pred. No. 78;
Matches 17; Conservative 0; Mismatches 1; Indels 1; Gaps 0;
```

```
QY      4  AGATAGGCACAAATGA 21
          ||||| ||||| ||||| |||||
Db      503  AGATAGGCACAAACAGA 520
```

RESULT 4

```
US-09-822-871-3/c
; Sequence 3, Application US/09822871
; Patent No. 6723547
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; AND USES THEREOF
; FILE REFERENCE: CL001219
; CURRENT APPLICATION NUMBER: US/09/822,871
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 4
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 254366
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(254366)
; OTHER INFORMATION: n = A,T,C or G
US-09-822-871-3
```

```
Query Match      78.1%; Score 16.4; DB 4; Length 254366;
Best Local Similarity 94.4%; Pred. No. 2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      4  AGATAGGCACAAATGA 21
          ||||| ||||| ||||| |||||
Db      228707 AGATATGCACAAATGA 228690
```

RESULT 5

```
US-09-949-016-12900/c
; Sequence 12900, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12900
; LENGTH: 462589
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12900
```

```
Query Match      78.1%; Score 16.4; DB 4; Length 462589;
Best Local Similarity 94.4%; Pred. No. 2.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY      4  AGATAGGCACAAATGA 21
          ||||| ||||| ||||| |||||
Db      101979 AGATAGGCACAAACAGA 101962
```

RESULT 6

```
US-09-949-016-12412/c
; Sequence 12412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```



```
; SEQ ID NO 12412
; LENGTH: 476044
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12412

Query Match          78.1%; Score 16.4; DB 4; Length 476044;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 AGATAGGCACCAACAATGA 21
      |||||
Db      115435 AGATAGGCACCAACAAGA 115418

RESULT 7
US-09-557-884-1
; Sequence 1, Application US/09557884
; Patent No. 6506581
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/557,884
; FILING DATE: 25-Apr-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/476,102
; FILING DATE: JUN-5-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Michelle S. Marks
; REGISTRATION NUMBER: 41,971
; REFERENCE/DOCKET NUMBER: PB186P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-557-884-1

Query Match          78.1%; Score 16.4; DB 4; Length 1830121;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 GAGATAGGCACCAACAATG 20
      |||||
Db      72931 GAGATAGGCACCAACAATG 72948

RESULT 8
US-09-643-990A-1
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5744
; LENGTH: 294
; TYPE: DNA
```

```
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; COMPUTER: Dell Pentium
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB186P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match          78.1%; Score 16.4; DB 4; Length 1830121;
Best Local Similarity 94.4%; Pred. No. 2.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 GAGATAGGCACCAACAATG 20
      |||||
Db      72931 GAGATAGGCACCAACAATG 72948

RESULT 9
US-09-313-294A-5744/c
; Sequence 5744, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 5744
; LENGTH: 294
; TYPE: DNA
```

```

; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700350613H1
US-09-3113-294A-5744

```

Query Match 77.1%; Score 16.2; DB 4; Length 294;
Best Local Similarity 85.7%; Pred. No. 88;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAACATGA 21
268 GTTAGATATGCACAACATTA 248
Db

```

RESULT 10
US-09-949-016-78348
; Sequence 78348, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78348
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-78348

```

```
Query Match      77.1%; Score 16.2; DB 4; Length 601;
Best Local Similarity 85.7%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY
Db

1 GTGAGATAGGCACAACAAATGA 21
91 GAGAGATAAGCACAATAATGA 111

```

RESULT 11
US-09-949-016-167629/c
; Sequence 167629, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 167629
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-167629

```

Query Match	77.1%	Score 16.2;	DB 4;	Length 601;
Best Local Similarity	85.7%	Pred. No. 99;		
Matches 18; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Qy 1 GTGAGATAGGCACAACAATGA 21
486 GTGACATAGGGACAACAATAA 466
Db

RESULT 12
US-09-949-016-167736/c
; Sequence 167736, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 167736
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-167736

Query Match 77.1%; Score 16.2; DB 4; Length 601;
Best Local Similarity 85.7%; Pred. No. 99;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GTGAGATAGGCACAACAATGA 21
||| ||| ||| ||| ||| ||| |||
Db 486 GTGACATAGGCACAACAATAA 466

```

RESULT 13
US-09-248-796A-11382
; Sequence 11382, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES FOR DIAGNOSTICS AND THERAPEUTICS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCES: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 11382
; LENGTH: 699
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-11382

```

```
Query Match      77.1%; Score 16.2; DB 4; Length 699;
Best Local Similarity 85.7%; Pred.No.1e+02;
Matches 18: Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 GTGAGATAGGCACAACAATGA 21
Db 15 GTGGGATAGGAAAAACAATGA 35

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11785
; LENGTH: 74096
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11785

Query Match 77.1%; Score 16.2; DB 4; Length 74096;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACAATGA 21
|||||
Db 38334 GTGAGATAGTAACAAAATGA 38354

RESULT 19
US-09-949-016-16239
; Sequence 16239, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16239
; LENGTH: 74097
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16239

Query Match 77.1%; Score 16.2; DB 4; Length 74097;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACAATGA 21
|||||
Db 38334 GTGAGATAGTAACAAAATGA 38354

RESULT 20
US-09-949-016-16466
; Sequence 16466, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16466
; LENGTH: 78157
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16466

Query Match 77.1%; Score 16.2; DB 4; Length 78157;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACAATGA 21
|||||
Db 64172 GTGACATAGGCACAAACAATAA 64192

RESULT 21
US-09-949-016-16467
; Sequence 16467, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16467
; LENGTH: 78157
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16467

Query Match 77.1%; Score 16.2; DB 4; Length 78157;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACAATGA 21
|||||
Db 64172 GTGACATAGGCACAAACAATAA 64192

RESULT 22
US-09-949-016-13203
; Sequence 13203, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13203
; LENGTH: 94019
; TYPE: DNA
; ORGANISM: Human
; FEATURE:

; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14033
; LENGTH: 784019
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(784019)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14033

Query Match 77.1%; Score 16.2; DB 4; Length 784019;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACCAACAATGA 21
| | | | | | | | | | | | | | | | | | | | | |
Db 68455 GAGAGATAAGCACCAATAATGA 68475

RESULT 27
US-09-949-016-12777
; Sequence 12777, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12777
; LENGTH: 828152
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(828152)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12777

Query Match 77.1%; Score 16.2; DB 4; Length 828152;
Best Local Similarity 85.7%; Pred. No. 2.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACCAACAATGA 21
| | | | | | | | | | | | | | | | | | | | | |
Db 64588 GAGAGATAAGCACCAATAATGA 64608

RESULT 28
US-09-949-016-76772
; Sequence 76772, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 76772
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-76772

Query Match 76.2%; Score 16; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ATAGGCACCAACAATGA 21
| | | | | | | | | | | | | | | | | | | | | |
Db 206 ATAGGCACCAACAATGA 221

RESULT 29
US-09-949-016-17000/c
; Sequence 17000, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17000
; LENGTH: 57638
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(57638)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17000

Query Match 76.2%; Score 16; DB 4; Length 57638;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GATAGGCACCAACAATG 20
| | | | | | | | | | | | | | | | | | | | | |
Db 29120 GATAGGCACCAACAATG 29105

RESULT 30
US-09-949-016-13969/c
; Sequence 13969, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13969
; LENGTH: 301828
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(301828)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13969

Query Match          76.2%; Score 16; DB 4; Length 301828;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 ATAGGCACACAAATGA 21
Db      284089 ATAGGCACACAAATGA 284074

RESULT 31
US-09-107-532A-3337/c
; Sequence 3337, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3337:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1497 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...1497
; SEQUENCE DESCRIPTION: SEQ ID NO: 3337:

US-09-107-532A-3337
Query Match          75.2%; Score 15.8; DB 4; Length 1497;
Best Local Similarity 89.5%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GTGAGATAGGCACAAAT 19
Db      584 GTGAGATAGGCACATAAT 566

RESULT 32
US-09-949-016-14249
; Sequence 14249, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14249
; LENGTH: 12995
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14249

Query Match          75.2%; Score 15.8; DB 4; Length 12995;
Best Local Similarity 89.5%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      3 GAGATAGGCACAAATGA 21
Db      3850 GTGAAAGGCACAAATGA 3868

RESULT 33
US-09-949-016-14250
; Sequence 14250, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14250
; LENGTH: 12995
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14250

Query Match          75.2%; Score 15.8; DB 4; Length 12995;
Best Local Similarity 89.5%; Pred. No. 2.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```


OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Query Match 75.2%; Score 15.8; DB 4; Length 1664976;
Best Local Similarity 89.5%; Pred.No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGAGATAGGCACACAATG 20
||||| ||||| ||||| ||||| |||||
Db 172246 TGAGAAAGACACACAATG 172264

Search completed: June 4, 2005, 11:53:30
Job time : 59.7219 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 07:08:50 ; Search time 200.183 Seconds
(without alignments)
644.888 Million cell updates/sec

Title: US-09-674-277-27
Perfect score: 21
Sequence: 1 gtagataggaacacaaatga 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5706582 seqs, 3073711274 residues

Total number of hits satisfying chosen parameters: 11413164

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.8	84.8	650	13	US-10-027-632-201759 Sequence 201759,
2	17.8	84.8	650	17	US-10-027-632-201759 Sequence 201759,
3	17.4	82.9	2643	17	US-10-108-260A-754 Sequence 754, Appl
4	16.8	80.0	240	9	US-09-974-300-4351 Sequence 4351, Ap
5	16.8	80.0	2566	10	US-09-798-889-34 Sequence 34, Appl
6	16.8	80.0	2566	17	US-10-633-680-34 Sequence 34, Appl
7	16.8	80.0	83709	18	US-10-723-860-2784 Sequence 2784, Ap
8	16.8	80.0	123192	14	US-10-175-523-71 Sequence 71, Appl
9	16.8	80.0	1503841	9	US-09-795-668-1 Sequence 1, Appli
10	16.8	80.0	1503841	9	US-09-795-668-1 Sequence 1, Appli
11	16.8	80.0	1503841	9	US-09-946-807-1 Sequence 1, Appli

C 12	16.8	80.0	1980090	18	US-10-719-993-6815 Sequence 6815, Ap
C 13	16.8	80.0	1980090	19	US-10-741-600-37676 Sequence 17676, A
C 14	16.4	78.1	25	19	US-10-719-900-839717 Sequence 839717, A
C 15	16.4	78.1	842	13	US-10-027-632-29251 Sequence 29251, A
C 16	16.4	78.1	842	13	US-10-027-632-29252 Sequence 29252, A
C 17	16.4	78.1	842	17	US-10-027-632-29251 Sequence 29251, A
C 18	16.4	78.1	842	17	US-10-027-632-29252 Sequence 29252, A
C 19	16.4	78.1	1710	13	US-10-027-632-99566 Sequence 99566, A
C 20	16.4	78.1	1710	17	US-10-027-632-99566 Sequence 99566, A
C 21	16.4	78.1	254366	10	US-09-822-871-3 Sequence 3, Appli
C 22	16.4	78.1	254366	17	US-10-673-885-3 Sequence 1, Appli
C 23	16.4	78.1	1830121	17	US-10-329-670-1 Sequence 1, Appli
C 24	16.4	78.1	1830121	18	US-10-158-865-1 Sequence 18, Appli
C 25	16.2	77.1	395	17	US-10-240-425-18 Sequence 72311, A
C 26	16.2	77.1	418	13	US-10-027-632-72311 Sequence 72311, A
C 27	16.2	77.1	418	17	US-10-027-632-72311 Sequence 73016, A
C 28	16.2	77.1	516	13	US-10-027-632-73016 Sequence 73016, A
C 29	16.2	77.1	516	17	US-10-027-632-73016 Sequence 73016, A
C 30	16.2	77.1	544	13	US-10-027-632-45756 Sequence 45756, A
C 31	16.2	77.1	544	17	US-10-027-632-45756 Sequence 45756, A
C 32	16.2	77.1	587	13	US-10-027-632-71212 Sequence 71212, A
C 33	16.2	77.1	587	17	US-10-027-632-71212 Sequence 71212, A
C 34	16.2	77.1	602	18	US-10-021-323-7671 Sequence 7671, Ap
C 35	16.2	77.1	624	18	US-10-021-323-8699 Sequence 8699, Ap
C 36	16.2	77.1	810	18	US-10-437-963-37663 Sequence 37663, A
C 37	16.2	77.1	1964	17	US-10-369-493-30459 Sequence 30459, A
C 38	16.2	77.1	4977	13	US-10-027-632-76128 Sequence 76128, A
C 39	16.2	77.1	4977	17	US-10-027-632-76128 Sequence 76128, A
C 40	16.2	77.1	44848	9	US-09-776-874A-42 Sequence 42, Appl
C 41	16.2	77.1	44848	9	US-09-988-113-42 Sequence 42, Appl
C 42	16.2	77.1	44848	16	US-10-341-582-42 Sequence 42, Appl
C 43	16.2	77.1	44848	16	US-10-384-451-42 Sequence 42, Appl
C 44	16.2	77.1	44848	16	US-10-384-450-42 Sequence 42, Appl
C 45	16.2	77.1	44848	17	US-10-371-219A-42 Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-10-027-632-201759
; Sequence 201759, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201759
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201759
Query Match 84.8%; Score 17.8; DB 13; Length 650;

Best Local Similarity 90.5%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAATGA 21
Db 457 GTGAGAAAGGCAAAACATGA 477

RESULT 2
US-10-027-632-201759
; Sequence 201759, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201759
; LENGTH: 650
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-201759

Query Match 84.8%; Score 17.8; DB 17; Length 650;
Best Local Similarity 90.5%; Pred. No. 83;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAATGA 21
Db 457 GTGAGAAAGGCAAAACATGA 477

RESULT 3
US-10-108-260A-754
; Sequence 754, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 754
; LENGTH: 2643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-754

Query Match 82.9%; Score 17.4; DB 17; Length 2643;
Best Local Similarity 94.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAGATAGGCACAAATGA 21

Db 1325 GAGATAGGCAGAACATGA 1343

RESULT 4
US-09-974-300-4351
; Sequence 4351, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; TITLE OF INVENTION: Expression
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4351
; LENGTH: 240
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-4351

Query Match 80.0%; Score 16.8; DB 9; Length 240;
Best Local Similarity 90.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TGAGATAGGCACAAATGA 21
Db 60 TGAGATGGGCACACATGA 79

RESULT 5
US-09-798-889-34/C
; Sequence 34, Application US/09798889
; Publication No. US20030004324A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 31 Human secreted proteins
; FILE REFERENCE: P2026P1
; CURRENT APPLICATION NUMBER: US/09/798,889
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/393,022
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-09
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,714
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,686
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,687
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/077,696
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 2566
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2553)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-798-889-34

Query Match 80.0%; Score 16.8; DB 10; Length 2566;
Best Local Similarity 90.0%; Pred. No. 3.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Query Match      80.0%; Score 16.8; DB 18; Length 83709;
Best Local Similarity 90.0%; Pred. No. 5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

77880 TGAGTAAAGCACAGCAATGA 77899

SSULT 8

S-10-175-523-71

Sequence 71, Application US/10175523

Publication No. US20030096264A1

GENERAL INFORMATION:

APPLICANT: Brockman, Jeffrey

APPLICANT: Evans, David

APPLICANT: Hook, Derek

APPLICANT: Klimczak, Leszek

APPLICANT: Laeng, Pascal

APPLICANT: Palfreyman, Michael

APPLICANT: Rajan, Prithi

TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)

FILE REFERENCE: 3235/IJ795-US3

CURRENT APPLICATION NUMBER: US/10/175,523

CURRENT FILING DATE: 2002-06-18

PRIOR APPLICATION NUMBER: US 60/299,151

PRIOR FILING DATE: 2001-06-18

PRIOR APPLICATION NUMBER: US 60/317,828

PRIOR FILING DATE: 2001-09-07

PRIOR APPLICATION NUMBER: US 60/325,150

PRIOR FILING DATE: 2001-09-25

PRIOR APPLICATION NUMBER: US 60/333,047

PRIOR FILING DATE: 2001-11-14

PRIOR APPLICATION NUMBER: US 60/349,936

PRIOR FILING DATE: 2002-01-18

PRIOR APPLICATION NUMBER: US 60/361,834

PRIOR FILING DATE: 2002-03-04

NUMBER OF SEQ ID NOS: 197

SOFTWARE: PatentIn version 3.1

SEQ ID NO 71

GENBANK: J25132
 TYPE: DNA
 ORGANISM: Mus musculus
 3-10-175-523-71

Query Match 80.0%; Score 16.8; DB 14; Length 123192;
Best Local Similarity 90.0%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

83833 TGAGAGAGGCACACAAAGA 83852
 5-09-795-668-1
 RESULT 9
 Sequence 1, Application US/09795668
 Patent No. US20020045577A1
 GENERAL INFORMATION:
 APPLICANT: Stefansson, Hreinn
 APPLICANT: Steinthoroddottir, Valgerdur
 APPLICANT: Gulcher, Jeffrey R.
 TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
 FILE REFERENCE: 2345.2004-001
 CURRENT APPLICATION NUMBER: US/09/795,668
 CURRENT FILING DATE: 2001-02-28
 PRIOR APPLICATION NUMBER: US 09/515,716
 PRIOR FILING DATE: 2000-02-28
 NUMBER OF SEQ ID NOS: 1531
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 1
 LENGTH: 1503841

```
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: m=a or a
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: y=t/u or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: m=a or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: k=g or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: s=g or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: w=a or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: b=g or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: s=g or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: d=a or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: h=a or c or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: v=a or g or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: n=a or g or c or t/u
/ OTHER INFORMATION: n=a or g or c or t/u
US-09-795-686-1

Query Match      80.0%; Score 16.8; DB 9; Length 1503841;
Best Local Similarity 90.0%; Pred. No. 6.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GTGAGTAGGCACCAATG 20
Db      1386660 GTGTGATAGGCACACACTG 1386679

RESULT 10
US-09-795-686-1
/ Sequence 1, Application US/09795686
/ Patent No. US20020094954A1
/ GENERAL INFORMATION:
/ APPLICANT: Stefansson, Hreinn
/ APPLICANT: Steinthorsdottir, Valgerdur
/ APPLICANT: Gulcher, Jeffrey R.
/ TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
/ FILE REFERENCE: 2345.2005-001
/ CURRENT APPLICATION NUMBER: US/09/795,686
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: US 09/515,715
/ PRIOR FILING DATE: 2000-02-28
/ NUMBER OF SEQ ID NOS: 1531
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1503841
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: r=g or a
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: y=t/u or c
```

```
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: m=a or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: k=g or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: s=g or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: w=a or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: b=g or c or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: d=a or g or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: h=a or c or t/u
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: v=a or g or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: n=a or g or c or t/u
/ OTHER INFORMATION: n=a or g or c or t/u
US-09-795-686-1

Query Match      80.0%; Score 16.8; DB 9; Length 1503841;
Best Local Similarity 90.0%; Pred. No. 6.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 GTGAGTAGGCACCAATG 20
Db      1386660 GTGTGATAGGCACACACTG 1386679

RESULT 11
US-09-946-807-1
/ Sequence 1, Application US/09946807
/ Patent No. US20020165144A1
/ GENERAL INFORMATION:
/ APPLICANT: Stefansson, Hreinn
/ APPLICANT: Steinthorsdottir, Valgerdur
/ APPLICANT: Gulcher, Jeffrey R.
/ TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
/ FILE REFERENCE: 2345.2004-001
/ CURRENT APPLICATION NUMBER: US/09/946,807
/ PRIOR FILING DATE: 2001-09-05
/ PRIOR APPLICATION NUMBER: US/09/795,668
/ PRIOR FILING DATE: 2001-02-28
/ PRIOR APPLICATION NUMBER: US 09/515,716
/ PRIOR FILING DATE: 2000-02-28
/ NUMBER OF SEQ ID NOS: 1531
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1503841
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: r=g or a
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: y=t/u or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: m=a or c
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(1531)
/ OTHER INFORMATION: y=t/u or c
```

NAME/KEY: misc feature
LOCATION: (1)...(1531)
OTHER INFORMATION: k=g or t/u
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1531)
OTHER INFORMATION: s=g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1531)
OTHER INFORMATION: w=a or t/u
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1531)
OTHER INFORMATION: b=g or c or t/u
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1531)
OTHER INFORMATION: d=a or g or t/u
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1531)
OTHER INFORMATION: h=a or c or t/u
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1531)
OTHER INFORMATION: v=a or g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(1531)
OTHER INFORMATION: n=a or g or c or t/u
US-09-946-807-1

Query Match 80.0%; Score 16.8; DB 9; Length 1503841;
Best Local Similarity 90.0%; Pred. No. 6.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GTGAGATAGGCACAACTG 20
Db 1386660 GTGTGATAGGCACAACTG 1386679

RESULT 12
US-10-719-993-6815/c
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6815

Query Match 80.0%; Score 16.8; DB 18; Length 1980090;
Best Local Similarity 90.0%; Pred. No. 6.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 TGAGATAGGCACAACTGA 21
Db 539439 TGAGAGGGCACAACTATGA 539420

RESULT 13
US-10-741-600-17676/c
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17676
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-741-600-17676

Query Match 80.0%; Score 16.8; DB 19; Length 1980090;
Best Local Similarity 90.0%; Pred. No. 6.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 TGAGATAGGCACAACTGA 21
Db 539439 TGAGAGGGCACAACTATGA 539420

RESULT 14
US-10-719-900-839717/c
; Sequence 839717, Application US/10719900
; Publication No. US20050026164A1
; GENERAL INFORMATION:
; APPLICANT: Xue Mei Zhou
; TITLE OF INVENTION: Methods of Genetic Analysis of Mouse
; FILE REFERENCE: 3528.1
; CURRENT APPLICATION NUMBER: US/10/719,900
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/427,808
; PRIOR FILING DATE: 2002 11 20
; NUMBER OF SEQ ID NOS: 982914
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 839717
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-719-900-839717

Query Match 78.1%; Score 16.4; DB 19; Length 25;
Best Local Similarity 94.4%; Pred. No. 2.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 AGATAGGCACAACTGA 21
Db 23 AGATTGGCACAACTGA 6

RESULT 15
US-10-027-632-29251/c
; Sequence 29251, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29251
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(842)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-29251

Query Match 78.1%; Score 16.4; DB 13; Length 842;

Best Local Similarity 94.4%; Pred. No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 AGATAGGCACACAAATGA 21
Db 70 AGATAGGCACATCAATGA 53

RESULT 16

US-10-027-632-29252/c
; Sequence 29252, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29252
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(842)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-29252

Query Match 78.1%; Score 16.4; DB 13; Length 842;

Best Local Similarity 94.4%; Pred. No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 4 AGATAGGCACACAAATGA 21
Db 70 AGATAGGCACATCAATGA 53

RESULT 17

US-10-027-632-29251/c
; Sequence 29251, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29251
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(842)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-29251

Query Match

78.1%; Score 16.4; DB 17; Length 842;

Best Local Similarity 94.4%; Pred. No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 AGATAGGCACACAAATGA 21
Db 70 AGATAGGCACATCAATGA 53

RESULT 18

US-10-027-632-29252/c
; Sequence 29252, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24

Query Match

; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29252
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(842)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-29252

Query Match 78.1%; Score 16.4; DB 17; Length 842;
Best Local Similarity 94.4%; Pred. No. 4.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AGATAGGCACACAAATGA 21
|||||
Db 70 AGATAGGCACATCAATGA 53

RESULT 19
US-10-027-632-99566
; Sequence 99566, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99566
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-99566

Query Match 78.1%; Score 16.4; DB 13; Length 1710;
Best Local Similarity 94.4%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AGATAGGCACACAAATGA 21
|||||
Db 1641 AGATAGGCACATCAATGA 1658

RESULT 20
US-10-027-632-99566
; Sequence 99566, Application US/10027632
; Publication No. US20030204075A9

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 99566
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-99566

Query Match 78.1%; Score 16.4; DB 17; Length 1710;
Best Local Similarity 94.4%; Pred. No. 4.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AGATAGGCACACAAATGA 21
|||||
Db 1641 AGATAGGCACATCAATGA 1658

RESULT 21
US-09-822-871-3/c
; Sequence 3, Application US/09822871
; Publication No. US20030099942A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS,
; FILE REFERENCE: CL001219
; CURRENT APPLICATION NUMBER: US/09/822,871
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 254366
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(254366)
; OTHER INFORMATION: n = A,T,C or G
US-09-822-871-3

Query Match 78.1%; Score 16.4; DB 10; Length 254366;
Best Local Similarity 94.4%; Pred. No. 8.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 AGATAGGCACACAAATGA 21
|||||
Db 228707 AGATAGGCACACAAATGA 228690

```
; Sequence 3, Application US/10673885
; Publication No. US20040081644A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHATASE PROTEINS.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHATASE PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001219DIV
; CURRENT APPLICATION NUMBER: US/10/673,885
; CURRENT FILING DATE: 2003-09-30
; PRIOR APPLICATION NUMBER: 09/822,871
; PRIOR FILING DATE: 04-02-2001
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 254366
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(254366)
; OTHER INFORMATION: n = A,T,C or G
US-10-673-885-3

Query Match      78.1%; Score 16.4; DB 17; Length 254366;
Best Local Similarity 94.4%; Pred. No. 8.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 AGATAGGCACACAAATGA 21
Db      228707 AGATATGCACACAAATGA 228690

RESULT 23
US-10-329-670-1
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; FILE REFERENCE: PB186P1
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29258)..(29298)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
```

; LOCATION: (10091)..(10091)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119750)..(119750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (119924)..(119924)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (120038)..(120038)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121344)..(121344)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122167)..(122167)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (122336)..(122336)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131340)..(131340)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (131360)..(131360)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (139910)..(139910)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (140398)..(140398)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (142750)..(142750)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145058)..(145058)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145171)..(145171)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (145942)..(145942)

; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (147197)..(147197)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Query Match 78.1%; Score 16.4; DB 17; Length 1830121;
Best Local Similarity 94.4%; Pred. No. 9.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAGATAGGCACCAATG 20
||| ||||| |||||
Db 72931 GAGATAGGCACCAATG 72948

RESULT 24

US-10-158-865-1
; Sequence 1, Application US/10158865
; Publication No. US20040203093A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fra
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g

FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:

NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature

```
; LOCATION: (150841)..(150841)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152500)..(152500)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (152530)..(152530)

Query Match      78.1%; Score 16.4; DB 18; Length 1830121;
Best Local Similarity 94.4%; Pred. No. 9.9e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      3 GAGATAGGCACACAATG 20
      |||||
Db      72931 GAGATAGGCACACAATG 72948

RESULT 25
US-10-240-425-18/c
; Sequence 18, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 18
; LENGTH: 395
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AA026238
US-10-240-425-18

Query Match      77.1%; Score 16.2; DB 17; Length 395;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GTGAGTAGGCACACAATGA 21
      |||||
Db      252 GTGAGTAGTAACAAAAATGA 232

RESULT 26
US-10-027-632-72311
; Sequence 72311, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72311
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72311

Query Match      77.1%; Score 16.2; DB 17; Length 418;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GTGAGTAGGCACACAATGA 21
      |||||
Db      99 GTGAGTAAGAGACACAATGA 119

RESULT 28
US-10-027-632-73016
; Sequence 73016, Application US/10027632
; Publication No. US20020198371A1
```

```
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72311
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72311

Query Match      77.1%; Score 16.2; DB 13; Length 418;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GTGAGTAGGCACACAATGA 21
      |||||
Db      99 GTGAGTAAGAGACACAATGA 119

RESULT 27
US-10-027-632-72311
; Sequence 72311, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72311
; LENGTH: 418
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72311

Query Match      77.1%; Score 16.2; DB 17; Length 418;
Best Local Similarity 85.7%; Pred. No. 4.9e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 GTGAGTAGGCACACAATGA 21
      |||||
Db      99 GTGAGTAAGAGACACAATGA 119

RESULT 28
US-10-027-632-73016
; Sequence 73016, Application US/10027632
; Publication No. US20020198371A1
```

GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73016
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-73016

Query Match 77.1%; Score 16.2; DB 13; Length 516;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACATGA 21
||||| ||| |||||
Db 99 GTGAGAAAGAGACACACATGA 119

RESULT 29
US-10-027-632-73016
; Sequence 73016, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 73016
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-73016

Query Match 77.1%; Score 16.2; DB 17; Length 516;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;

Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 GTGAGATAGGCACAAACATGA 21
||||| ||| |||||
Db 99 GTGAGAAAGAGACACACATGA 119
RESULT 30
US-10-027-632-45756
; Sequence 45756, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45756
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-45756

Query Match 77.1%; Score 16.2; DB 13; Length 544;
Best Local Similarity 85.7%; Pred. No. 5.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACATGA 21
||||| ||| |||||
Db 115 GTGAGATAGACACACACTAA 135

RESULT 31
US-10-027-632-45756
; Sequence 45756, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

; Sequence 8699, Application US/10021323
; Publication No. US20040123340A1
; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Peng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)B
; CURRENT APPLICATION NUMBER: US/10/021,323
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 60/255, 619
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 8699
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(624)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3828-019-Q1-K6-A9
US-10-021-323-8699

Query Match 77.1%; Score 16.2; DB 18; Length 624;
Best Local Similarity 85.7%; Pred. No. 5.2e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAACTGA 21
|||||
Db 421 GTGACATAGCACATCAATGA 401

RESULT 36

US-10-437-963-37663
; Sequence 37663, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 37663
; LENGTH: 810
; TYPE: DNA

; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_41372C.1
US-10-437-963-37663

Query Match 77.1%; Score 16.2; DB 18; Length 810;
Best Local Similarity 85.7%; Pred. No. 5.4e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAACTGA 21
|||||
Db 716 GTGAGGAAGGACAACTGA 736

RESULT 37

US-10-369-493-30459/c

; Sequence 30459, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 30459
; LENGTH: 1964
; TYPE: DNA

; ORGANISM: Caenorhabditis elegans
US-10-369-493-30459

Query Match 77.1%; Score 16.2; DB 17; Length 1964;
Best Local Similarity 85.7%; Pred. No. 6e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAACTGA 21
|||||
Db 1928 GTGACATAGGCATACTATGA 1908

RESULT 38

US-10-027-632-76128/c

; Sequence 76128, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 76128
; LENGTH: 4977
; TYPE: DNA

; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature

; LOCATION: (1)_(4977)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-76128

Query Match 77.1%; Score 16.2; DB 13; Length 4977;
Best Local Similarity 85.7%; Pred. No. 6.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAACTGA 21


```
Db      4854 GTGGGAGGACACAAATGA 4834
|||||
RESULT 39
US-10-027-632-76128/c
; Sequence 76128, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 10827-129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76128
; LENGTH: 4977
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(4977)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-76128

Query Match      77.1%; Score 16.2; DB 17; Length 4977;
Best Local Similarity 85.7%; Pred. No. 6.8e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 GTGAGATAGGCACAAATGA 21
|||||
Db      4854 GTGGGAGGACACAAATGA 4834

RESULT 40
US-09-776-874A-42/c
; Sequence 42, Application US/09776874A
; Patent No. US20020102560A1
; GENERAL INFORMATION:
; APPLICANT: Pecker, Iris
; APPLICANT: Vlodevsky, Israel
; APPLICANT: Feinstein, Elena
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
; EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
; FILE REFERENCE: 01/22603
; CURRENT APPLICATION NUMBER: US/09/776,874A
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: US 08/922,170
; PRIOR FILING DATE: 1997-09-02
; PRIOR APPLICATION NUMBER: US 09/109,386
; PRIOR FILING DATE: 1998-07-10
; PRIOR APPLICATION NUMBER: PCT/US98/17954
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 44848
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-776-874A-42

Query Match      77.1%; Score 16.2; DB 9; Length 44848;
Best Local Similarity 85.7%; Pred. No. 9.1e+02;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 GTGAGATAGGCACAAATGA 21
|||||
Db      24187 GTTAGATTGGCACCACAATGA 24167

Search completed: June 4, 2005, 12:20:31
Job time : 215.183 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2005, 06:32:00 ; Search time 1282.12 Seconds
(without alignments)
623.460 Million cell updates/sec

Title: US-09-674-277-27
Perfect score: 21
Sequence: 1 gtgagataggcaacaatga 21

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gss1:*
- 9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19.4	92.4	581	1	AL699132 DKFZp686D
C 2	18.4	87.6	219	2	BF761423 RC0-CS004
C 3	18	85.7	567	8	AZ245932 RPCI-23-1
C 4	17.8	84.8	411	1	AI204089 QD73C03.X
C 5	17.8	84.8	440	8	AZ787729 2M0034D22
C 6	17.8	84.8	473	2	AW272008 xs12h11.x
C 7	17.8	84.8	660	9	CE588178 tigr-gss-
C 8	17.8	84.8	665	9	CE699569 tigr-gss-
C 9	17.8	84.8	679	9	CE176088 tigr-gss-
C 10	17.8	84.8	704	4	BJ709508 BJ709508
C 11	17.8	84.8	801	7	CK148092
C 12	17.8	84.8	802	7	CK127649 AGENCOURT
C 13	17.8	84.8	811	7	CK140761
C 14	17.8	84.8	833	7	CK141470 AGENCOURT
C 15	17.8	84.8	851	7	CK143478 AGENCOURT
C 16	17.8	84.8	857	6	CD756258 AGENCOURT
C 17	17.8	84.8	860	6	CD755916 AGENCOURT
C 18	17.8	84.8	928	9	CL470941 SAIL_151
C 19	17.8	84.8	1500	1	AL931497 AL931497
C 20	17.4	82.9	306	1	AV045048 AV045048
C 21	17.4	82.9	319	6	CAS18688 K310012E0
C 22	17.4	82.9	357	8	AQ242274 HS 2058.B
C 23	17.4	82.9	556	7	CF937946 NCSTqab5
C 24	17.4	82.9	585	7	CF796899 NCSTqab8

C 25	17.4	82.9	600	7	CF423038
C 26	17.4	82.9	602	7	CF416715
C 27	17.4	82.9	717	5	BU093879 LX_CALZ_0
C 28	17.4	82.9	904	8	BZ119502 CH230-411
C 29	17.4	82.9	919	9	CNS048SL Tetraodon
C 30	17.4	82.9	946	8	BZ976367 PUDS91TD
C 31	17	81.0	743	7	CK173441 EST762761
C 32	17	81.0	743	7	CK173442 EST762762
C 33	17	81.0	895	9	CG951752 MBEJC69TF
C 34	16.8	80.0	158	8	AZ058773 RPCI-23-4
C 35	16.8	80.0	415	5	BY006779 BY006779
C 36	16.8	80.0	456	8	AQ338757 HS 2210.B
C 37	16.8	80.0	458	8	AQ819355 HS_5295_A
C 38	16.8	80.0	461	8	AQ803506 HS_2126_A
C 39	16.8	80.0	475	5	BU733937 UI-E-CKI-
C 40	16.8	80.0	499	8	AZ016072 RPCI-23-2
C 41	16.8	80.0	513	9	CL624963 OR_BBA002
C 42	16.8	80.0	524	2	BF558492 UI-R-C1-1
C 43	16.8	80.0	533	8	AZ057095 RPCI-23-4
C 44	16.8	80.0	556	8	AZ252122 RPCI-23-4
C 45	16.8	80.0	571	9	CE238399 tigr-gss-

ALIGNMENTS

RESULT 1
AL699132/c
LOCUS
DEFINITION DKFZp686D16112 r1.686 (synonym: hlccc3) Homo sapiens CDNA clone
ACCESSION AL699132 581 bp mRNA linear EST 04-SEP-2003
VERSION DKFZp686D16112.5, mRNA sequence.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS EST (bases 1 to 581)
TITLES Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Poustka,A., Wellenreuther,R., Mewes,H.W., Weil,B. and Wiemann,S.).
JOURNAL Wiemann,S.)
COMMENT Unpublished (1999)
Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the CDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp686D16112) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
1..581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686D16112"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlccc3)"
/note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 92.4%; Score 19.4; DB 1; Length 581;
Best Local Similarity 95.2%; Pred. No. 79;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
 ||||| ||||| ||||| |||||
 Db 263 GTGAGAAAGGCACACAAATGA 243

RESULT 2

BF761423/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

219 bp mRNA linear EST 12-JAN-2001
 RC0-CS0045-041000-021-a12 CS0045 Homo sapiens cDNA, mRNA sequence.
 BF761423
 BF761423
 BF761423.1 GI:12109323
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 219)
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=RC0&ct=RC0-CS0045-
 041000-021-a12&t3=2000-10-04&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 21
 High quality sequence stop: 219.

FEATURES

source

1..219
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CS0045"
 /note="Organ: colon; est; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 87.6%; Score 18.4; DB 2; Length 219;
 Best Local Similarity 95.0%; Pred. No. 2.3e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TCGATAGGCACACAAATGA 21

||| ||||| ||||| |||||

Db 103 TCGATAGGCACACAAATGA 84

RESULT 3

AZ245932/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

567 bp DNA linear GSS 15-JUN-2000
 RPCI-23-11N12.TJB RPCI-23 Mus musculus genomic clone
 Genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AZ245932

AZ245932.1 GI:8559128

GSS.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,

Akiret, B., Levine, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de

Jong, P. and Fraser, C.M.

Mouse BAC End Sequences from Library RPCI-23

Unpublished (1999)

Other GSSs: RPCI-23-11N12.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)or from Resea ch Genetics (info@resgen.com). BAC end page:http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html

Plate: 11 row: N column: 12

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..567

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="CS7BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-23-11N12"

/sex="Female"

/lab_host="DH10B"

/clone_lib="RPCI-23"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1:

EcoRI; Site 2: EcoRI; Female CS7BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested

with a combination of EcoRI and EcoRI Methylase. Size

selected DNA was cloned into the pBAC3.6 vector at the

EcoRI sites. The ligation products were transformed into

DH10B electrocompetent cells (BRL Life Technologies). "

ORIGIN

Query Match 85.7%; Score 18; DB 8; Length 567;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGATAGGCACACAA 18
 ||||| ||||| ||||| |||||
 Db 179 GTGATAGGCACACAA 162

RESULT 4

AI204089

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

411 bp mRNA linear EST 28-OCT-1998
 qd73c03.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1735108
 3', mRNA sequence.
 AI204089
 AI204089.1 GI:3756695
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 411)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),


```

FEATURES
source
High quality sequence stop: 447.
Location/Qualifiers
1. .473
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2769477"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid11"
/notes="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(clone1Ds 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."

ORIGIN
Query Match 84.8%; Score 17.8; DB 2; Length 473;
Best Local Similarity 90.5%; Pred. No. 4.9e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAATGA 21
Db 108 GTGAGGTAGGCACACAATTA 128

RESULT 7
CE588178 660 bp DNA linear GSS 28-SBP-2003
LOCUS tigr-gss-dog-17000366360022 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE588178.1 GI:36904959
VERSION GSS.
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 660)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness,E.F., Batna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .660
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 84.8%; Score 17.8; DB 9; Length 660;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAATGA 21
Db 508 GTGTGATAGGCACACAATGA 528

RESULT 9
CE176088 679 bp DNA linear GSS 25-SBP-2003
LOCUS tigr-gss-dog-17000326728063 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE176088.1 GI:35321060
VERSION GSS.
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 679)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness,E.F., Batna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .660
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 84.8%; Score 17.8; DB 9; Length 660;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 GTGAGATAGGCACACAATGA 21
Db 503 GTGTGATAGGCACACAATGA 523

RESULT 8
CE699569 665 bp DNA linear GSS 29-SBP-2003
LOCUS tigr-gss-dog-17000368812318 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE699569.1 GI:37018793
VERSION GSS.
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 665)
AUTHORS Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness,E.F., Batna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .665
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 84.8%; Score 17.8; DB 9; Length 665;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAATGA 21
Db 508 GTGTGATAGGCACACAATGA 528

RESULT 9
CE176088 679 bp DNA linear GSS 25-SBP-2003
LOCUS tigr-gss-dog-17000326728063 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE176088.1 GI:35321060
VERSION GSS.
KEYWORDS Canis familiaris (dog)
SOURCE Canis familiaris
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 679)
AUTHORS Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness,E.F., Batna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .665
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 84.8%; Score 17.8; DB 9; Length 665;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

COMMENT

Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

FEATURES
source

Location/Qualifiers
1..679
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 679;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACATGA 21
|||||
Db 259 GTGTATAGGTACACAAACATGA 279

RESULT 10
BJ709508/c

LOCUS 704 bp mRNA linear EST 08-MAR-2004
DEFINITION BJ709508 MF01FFA cDNA Oryzias latipes CDNA clone MF01FFA009g14 5',
mRNA sequence.

ACCESSION BJ709508
VERSION BJ709508.1 GI:45250452

KEYWORDS
SOURCE

Oryzias latipes (Japanese medaka)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.

REFERENCE

1 (bases 1 to 704)
Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
Medaka EST Project in Takeda's lab

TITLE

Unpublished (2001)

JOURNAL

Contact: Tadasu Shin-i
Center for Genetic Resource Information
National Institute of Genetics

COMMENT

1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855

FEATURES

Location/Qualifiers
1..704
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01FFA009g14"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="fry stage 40"
/clone_lib="MF01FFA CDNA"

ORIGIN

Query Match 84.8%; Score 17.8; DB 4; Length 704;
Best Local Similarity 90.5%; Pred. No. 5.1e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAACATGA 21
|||||
Db 411 GTGTATAGGTACACAAACATGA 391

RESULT 11
CK148092

LOCUS 801 bp mRNA linear EST 04-DEC-2003
DEFINITION CK148092 NCI CGAP ZEmb2 Danio rerio CDNA clone
IMAGE:7061234 5', mRNA sequence.

ACCESSION CK148092
VERSION CK148092.1 GI:38686717

KEYWORDS
SOURCE

Danio rerio (zebrafish)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 801)
NIH-MGC http://mgs.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished (1999)

JOURNAL

Contact: Daniela S. Gerhard, Ph.D.

COMMENT

Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Place: LLM1489 row: 9 column: 24

High quality sequence stop: 553.

FEATURES

Location/Qualifiers
1..801
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7061234"
/tissue_type="embryo"
/lab_host="DH10B (Tl-resistant)"
/clone_lib="NCI CGAP ZEmb2"
/note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI CGAP Library."

ORIGIN

Query Match 84.8%; Score 17.8; DB 7; Length 801;
Best Local Similarity 90.5%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY

1 GTGAGATAGGCACAAACATGA 21

Db

777 GTGAGATAGGCACAAACATGA 797

RESULT 12

LOCUS 802 bp mRNA linear EST 02-DEC-2003
DEFINITION CK127649 AGENCOURT_16705073 NIH_ZGC_10 Danio rerio CDNA clone IMAGE:7047179
5', mRNA sequence.

ACCESSION

CK127649

VERSION

CK127649.1 GI:38617095

KEYWORDS

Danio rerio (zebrafish)

SOURCE

Danio rerio

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE

1 (bases 1 to 802)
NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 National Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Len Zon, Harvard
 cDNA Library Preparation: Open Biosystems
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM14812 row: n column: 09
 High quality sequence stop: 676.

FEATURES
 source
 1..802
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7047179"
 /tissue_type="whole body"
 /lab_host="DH10B"
 /clone_lib="NIH_ZGC_10"
 /note="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
 Bulk tissue was collected from a whole adult individual
 from the Tuebingen strain. 1st strand cDNA was primed with
 a Not I - oligo(dT) primer, double-stranded cDNA was
 cloned into the Not I and EcoRV sites of pExpress-1.
 Library was size-selected for >1 kb fragments. A
 normalized version of this library is also available
 (NIH ZGC 7). Library was constructed by Open Biosystems
 (Huntsville, AL)."

ORIGIN
 Query Match 84.8%; Score 17.8; DB 7; Length 802;
 Best Local Similarity 90.5%; Pred. No. 5.2e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAACTGA 21
 |||||
Db 764 GTGAGACAGGCACAACTGA 784

RESULT 13
CK140761
LOCUS CK140761 811 bp mRNA linear EST 03-DEC-2003
DEFINITION AGENCOURT_16886306 NCI_CGAP_Zemb3 Danio rerio cDNA clone
ACCESSION IMAGE:7058518 5', mRNA sequence.
VERSION CK140761
KEYWORDS EST.
SOURCE CK140761.1 GI:38651687
ORGANISM Danio rerio (zebrafish)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 811)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Chi-Bin Chien
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

FEATURES
 source
 1..833
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7058635"
 /tissue_type="embryo"
 /lab_host="DH10B (T1-resistant)"
 /clone_lib="NCI CGAP_Zemb3"
 /note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
 Cloned unidirectionally. Primer: Oligo dT. Average insert
 size 2.1 kb. Constructed by J. Wang (Research Genetics,
 Invitrogen Corp) from tissue donated by L. Zon (Harvard
 University). Note: this is a NCI_CGAP Library."

http://image.llnl.gov
 Plate: LLAM14842 row: f column: 20
 High quality sequence start: 3
 High quality sequence stop: 745.

FEATURES
 source
 1..811
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7058518"
 /tissue_type="embryo"
 /lab_host="DH10B (T1-resistant)"
 /clone_lib="NCI CGAP_Zemb3"
 /note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
 Cloned unidirectionally. Primer: Oligo dT. Average insert
 size 2.1 kb. Constructed by J. Wang (Research Genetics,
 Invitrogen Corp) from tissue donated by L. Zon (Harvard
 University). Note: this is a NCI_CGAP Library."

ORIGIN
 Query Match 84.8%; Score 17.8; DB 7; Length 811;
 Best Local Similarity 90.5%; Pred. No. 5.2e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAACTGA 21
 |||||
Db 789 GTGACAGGCACAACTGA 809

RESULT 14
CK141470
LOCUS CK141470 833 bp mRNA linear EST 03-DEC-2003
DEFINITION AGENCOURT_16876821 NCI_CGAP_Zemb3 Danio rerio cDNA clone
ACCESSION IMAGE:7056635 5', mRNA sequence.
VERSION CK141470
KEYWORDS EST.
SOURCE CK141470.1 GI:38652396
ORGANISM Danio rerio (zebrafish)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 833)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Chi-Bin Chien
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM14837 row: h column: 09
 High quality sequence stop: 670.

FEATURES
 source
 1..833
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:7056635"
 /tissue_type="embryo"
 /lab_host="DH10B (T1-resistant)"
 /clone_lib="NCI CGAP_Zemb3"
 /note="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
 Cloned unidirectionally. Primer: Oligo dT. Average insert
 size 2.1 kb. Constructed by J. Wang (Research Genetics,
 Invitrogen Corp) from tissue donated by L. Zon (Harvard
 University). Note: this is a NCI_CGAP Library."

University). Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 84.8%; Score 17.8; DB 7; Length 833;
Best Local Similarity 90.5%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
|||||
Db 740 GTCAGACAGGCACACAAATGA 760

RESULT 15
CK143478 851 bp mRNA linear EST 03-DEC-2003
LOCUS
DEFINITION
AGENCOURT_16870436 NCI_CGAP_Zemb3 Danio rerio cDNA clone
IMAGE:7056956 5', mRNA sequence.

ACCESSION CK143478
VERSION CK143478
KEYWORDS CK143478.1 GI:38652596

SOURCE EST.
ORGANISM Danio rerio (zebrafish)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 851)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Chi-Bin Chien
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLMI4838 row: e column: 18
High quality sequence start: 2
High quality sequence stop: 736.

FEATURES
source
1..851
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7056956"
/tissue_type="embryo"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_Zemb3"
/notes="Vector: PCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.1 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

Query Match 84.8%; Score 17.8; DB 7; Length 851;
Best Local Similarity 90.5%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
|||||

Db 803 GTCAGACAGGCACACAAATGA 823

RESULT 16
CD756258 857 bp mRNA linear EST 30-JUN-2003
LOCUS
DEFINITION
AGENCOURT_14622483 NCI_CGAP_Zemb3 Danio rerio cDNA clone

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 857)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892

Tissue Procurement: Dr. Chi-Bin Chien
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLMI4838 row: e column: 18
High quality sequence start: 2
High quality sequence stop: 736.

FEATURES
source
1..851
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7056956"
/tissue_type="embryo"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_Zemb3"
/notes="Vector: PCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.1 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

IMAGE:6969665 5', mRNA sequence.
CD756258
CD756258.1 GI:32340545
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 857)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Chi-Bin Chien
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLMI4617 row: p column: 16
High quality sequence stop: 619.

FEATURES
source
1..857
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6969665"
/tissue_type="embryo"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_Zemb3"
/notes="Vector: PCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.1 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

Query Match 84.8%; Score 17.8; DB 6; Length 857;
Best Local Similarity 90.5%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
|||||

Db 764 GTCAGACAGGCACACAAATGA 784

RESULT 17
CD755916
LOCUS
DEFINITION
AGENCOURT_14618149 NCI_CGAP_Zemb3 Danio rerio cDNA clone

IMAGE:6972022 5', mRNA sequence.
CD755916
CD755916.1 GI:32340203
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 860)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892

Tissue Procurement: Dr. Chi-Bin Chien
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLMI4617 row: p column: 16
High quality sequence stop: 619.

FEATURES
source
1..857
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6969665"
/tissue_type="embryo"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_Zemb3"
/notes="Vector: PCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.1 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

Query Match 84.8%; Score 17.8; DB 6; Length 857;
Best Local Similarity 90.5%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
|||||

Db 764 GTCAGACAGGCACACAAATGA 784

RESULT 17
CD755916
LOCUS
DEFINITION
AGENCOURT_14618149 NCI_CGAP_Zemb3 Danio rerio cDNA clone
IMAGE:6972022 5', mRNA sequence.
CD755916
CD755916.1 GI:32340203
EST.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 860)

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892

Tissue Procurement: Dr. Chi-Bin Chien
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLMI4617 row: p column: 16
High quality sequence stop: 619.

FEATURES
source
1..857
Location/Qualifiers
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:6969665"
/tissue_type="embryo"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_Zemb3"
/notes="Vector: PCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.1 kb. Constructed by J. Wang (Research Genetics,
Invitrogen Corp) from tissue donated by L. Zon (Harvard
University). Note: this is a NCI_CGAP Library."

Query Match 84.8%; Score 17.8; DB 6; Length 857;
Best Local Similarity 90.5%; Pred. No. 5.2e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21
|||||

Db 764 GTCAGACAGGCACACAAATGA 784

Email: c9apbs-x@mail.nih.gov
 Tissue Procurement: Dr. Chi-Bin Chien
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM14623 row: b column: 21
 High quality sequence start: 9
 High quality sequence stop: 652.
 Location/Qualifiers

FEATURES

source

1..860
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="IMAGE:6972022"
 /tissue_type="embryo"
 /lab_host="DH10B (TI-resistant)"
 /clone_lib="NCI CGAP ZEmb3"
 /notes="Vector: pCMV-SPORT6.1; Site 1: EcoRV; Site 2: NotI;
 Cloned unidirectionally. Primer: Oligo dT. Average insert
 size 2.1 kb. Constructed by J. Wang (Research Genetics,
 Invitrogen Corp) from tissue donated by L. Zon (Harvard
 University). Note: this is a NCI CGAP Library."

ORIGIN

Query Match 84.8%; Score 17.8; DB 6; Length 860;
 Best Local Similarity 90.5%; Pred. No. 5.2e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAATGA 21
 |||||
 Db 796 GTCAGACAGGCACAAATGA 816

RESULT 18

CL470941/c

LOCUS

DEFINITION SAIL_151_E05.v1 SAIL Collection Arabidopsis thaliana genomic clone
 SAIL_151_E05.v1, genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

1 (bases 1 to 928)

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D.,
 Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J.D., Cotton, D.,
 Bullis, D., Snell, J., Miquel, T., Hutchison, D., Kimmerly, B.,
 Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S.A.
 A high-throughput Arabidopsis reverse genetics system
 Plant Cell 14 (12), 2985-2994 (2002)

PUBMED

CONTACT: Sessions A

Applied Trait Genetics

Syngenta Biotechnology Inc.

3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA

Email: allen.sessions@syngenta.com

ABRC Stock Number CS807328; T-DNA left border flanking sequences of
 Syngenta Arabidopsis Insertion Library (SAIL) lines are available
 through the Arabidopsis Biological Resource Center (ABRC).
 Sequences represent a pool of amplified genomic regions and not
 single contiguous sequences.
 Class: TDNA tagged.

Location/Qualifiers

1..928

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

FEATURES

source

/ecotype="Columbia"
 /db_xref="taxon:3702"
 /clone="SAIL 151_E05.v1"
 /clone_lib="SAIL_Collection"
 /note="T-DNA left border sequences were isolated using a
 modified TAIL-PCR strategy"

ORIGIN

Query Match 84.8%; Score 17.8; DB 9; Length 928;
 Best Local Similarity 90.5%; Pred. No. 5.3e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAATGA 21
 |||||
 Db 550 GTGAGAGAGGCAGAACAAATGA 530

RESULT 19

AL931497

LOCUS

DEFINITION AL931497 NAP1 Anopheles gambiae cDNA clone NAP1-P81-C-05-5, mRNA
 sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Anopheles gambiae (African malaria mosquito)

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
 Anopheles.

1 (bases 1 to 1500)

Christophides, G.K., Blass, K., Zdobnov, E.M., Carmouche, R., Benes, V.
 and Kafatos, F.C.
 Anopheles gambiae EST, European Molecular Biology Laboratory
 Unpublished (2002)
 Contact: Christophides GK
 Fotis C. Kafatos Laboratory
 European Molecular Biology Laboratory
 Meyerhofstrasse 1, 69117 Heidelberg, Germany
 Tel: +49 6221 387-440
 Fax: +49 6221 387-306
 Email: christop@embl-heidelberg.de
 Plate: P81 row: C column: 05.
 Location/Qualifiers

1..1500

/organism="Anopheles gambiae"

/mol_type="mRNA"

/db_xref="taxon:7165"

/clone="NAP1-P81-C-05-5"

/lab_host="E. coli DH10B"

/clone_lib="NAP1"

/note="Vector: pT7T3D-Pac (Pharmacia); Site 1: NotI;
 Site 2: EcoRI; ESTs sequenced from the T7 priming site
 that reads from the 5' end of cDNA. The NAP1 is a
 directionally cloned and normalized, oligo-T primed cDNA
 library constructed from a mixture of Anopheles gambiae
 developmental stages according to: Bonaldo, Lennon &
 Soares (1996): Normalization and Subtraction: Two
 Approaches To Facilitate Gene Discovery, Genome Research
 6, 791-806."

6, 791-806."

FEATURES

source

1..1500

/organism="Anopheles gambiae"

/mol_type="mRNA"

/db_xref="taxon:7165"

/clone="NAP1-P81-C-05-5"

/lab_host="E. coli DH10B"

/clone_lib="NAP1"

/note="Vector: pT7T3D-Pac (Pharmacia); Site 1: NotI;
 Site 2: EcoRI; ESTs sequenced from the T7 priming site
 that reads from the 5' end of cDNA. The NAP1 is a
 directionally cloned and normalized, oligo-T primed cDNA
 library constructed from a mixture of Anopheles gambiae
 developmental stages according to: Bonaldo, Lennon &
 Soares (1996): Normalization and Subtraction: Two
 Approaches To Facilitate Gene Discovery, Genome Research
 6, 791-806."

6, 791-806."

ORIGIN

Query Match 84.8%; Score 17.8; DB 1; Length 1500;
 Best Local Similarity 90.5%; Pred. No. 5.5e+02;
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACAAATGA 21
 |||||
 Db 174 GTGAGCTAGACACAAATGA 194

RESULT 20

AV045048/c

LOCUS AV045048 306 bp mRNA linear EST 23-NOV-1999
 DEFINITION AV045048 Mus musculus adult C57BL/6J testis Mus musculus cDNA clone
 1700041B03, mRNA sequence.
 ACCESSION AV045048
 VERSION AV045048.2 GI:4864713
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 306)
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 KARNINCI, P., SHIBATA, K., OZAWA, Y., KONNO, H., ITOH, M., AIZAWA, K., AKAHIRA, S., AKIYAMA, J., FUKUDA, S., FUKUNISHI, Y., FUNAYAMA, T., HARA, A., HAYATSU, N., HORI, F., ISHIKAWA, T., ITOH, M., IZAWA, M., KAWAI, J., KIKUCHI, N., KOJIMA, Y., MATSUYAMA, T., NIITSUMA, H., ODA, H., OWA, C., SATO, K., SHIBATA, Y., SHIGEMOTO, Y., SHIRAKI, T., SOGABE, Y., SUGAHARA, Y., SUZUKI, H., SUZUKI, H., TATENO, M., TOMARU, Y., TOMINAGA, N., WATANABE, S., YAGAME, M., YAMAMURA, T., YOKOTA, T., YOSHINO, M., MURAMATSU, M., OKAZAKI, Y. and HAYASHIZAKI, Y.
 TITLE RIKEN Mouse ESTs
 JOURNAL Unpublished (1999)
 COMMENT Contact: Chile Owa
 Genome Science Laboratory
 RIKEN
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-9145
 Fax: 81-298-36-9098
 Email: genome-res@rtc.riken.go.jp
 Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA (Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998))
 Transcriptional sequencing: A method for DNA sequencing using RNA polymerase (Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998))
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.
 FEATURES
 source Location/Qualifiers
 1..306
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="1700041B03"
 /sex="male"
 /tissue_type="testis"
 /dev_stage="adult"
 /clone_lib="Mus musculus adult C57BL/6J testis"
 ORIGIN
 Query Match 82.9%; Score 17.4; DB 1; Length 306;
 Best Local Similarity 94.7%; Pred. No. 7.5e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 GAGATAGGCACAAATGA 21
 |||||
 Db 24 GAATAGGCACAAATGA 6
 |||||
 RESULT 21
 CA518688/c 319 bp mRNA linear EST 15-NOV-2002
 LOCUS CA518688
 DEFINITION KS10012E05 KS10 Capsicum annuum cDNA, mRNA sequence.
 ACCESSION CA518688
 VERSION CA518688.1 GI:25019245
 KEYWORDS EST.
 SOURCE Capsicum annuum
 ORGANISM Capsicum annuum
 REFERENCE 1 (bases 1 to 319)
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanales; Solanaceae; Capsicum.
 Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.
 TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum

JOURNAL Response Against Pathogen
 COMMENT Contact: Doil Choi
 Genome Research Center and National Center for Genome Information
 Korea Research Institute of Bioscience and Biotechnology
 P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
 Tel: 82-42-860-4340
 Fax: 82-42-860-4309
 Email: doil@mail.kribb.re.kr
 Plate: 012 row: E column: 05.
 FEATURES
 source Location/Qualifiers
 1..319
 /organism="Capsicum annuum"
 /mol_type="mRNA"
 /db_xref="taxon:4072"
 /clone_lib="KS10"
 ORIGIN
 Query Match 82.9%; Score 17.4; DB 6; Length 319;
 Best Local Similarity 94.7%; Pred. No. 7.6e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 GAGATAGGCACAAATGA 21
 |||||
 Db 289 GAGATAGGCACAAATGA 271
 |||||
 RESULT 22
 AQ242274 357 bp DNA linear GSS 03-OCT-1998
 LOCUS HS_2058_B2_H05_MR CIT Approved Human Genomic Sperm Library D Homo
 DEFINITION sapiens genomic clone Plate=2058 Col=10 Row=P, genomic survey
 sequence.
 ACCESSION AQ242274
 VERSION AQ242274.1 GI:3688896
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 357)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Sequence Tagged Connector
 Plate: 2058 row: P column: 10
 Class: BAC ends
 High quality sequence stop: 357.
 FEATURES
 source Location/Qualifiers
 1..357
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="plate=2058 Col=10 Row=P"
 /sex="male"
 /notes="Organ: sperm; Vector: pBelBAC11; BAC Clones in E-Coli DH10B"
 ORIGIN

Query Match	82.9%; Score 17.4; DB 8; Length 357;	
Best Local Similarity	94.7%; Pred. No. 7.6e+02;	
Matches	18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 GTGAGATAGGCACACAAAT 19	
DB	192 GTGAGATAGGCACAAACAAAT 210	
RESULT 23		
CF937946/c		
LOCUS	556 bp mRNA linear EST 19-NOV-2003	
DEFINITION	NCST3qb27e08.y1 Nc-1 Tachyzoite cDNA Library 2 Neospora caninum	
ACCESSION	CDNA 5', mRNA sequence.	
VERSION	CF937946	
KEYWORDS	EST.	
SOURCE	CF937946.1 GI:38435745	
ORGANISM	Neospora caninum	
REFERENCE	1 (bases 1 to 556)	
AUTHORS	Cole, R., Fogarty, S., Tang, K., Howe, D. K., Sibley, L. D., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Ronko, I., Tsagareishvili, R., Fedele, M., Belaygorod, L., Franklin, C., Carr, L. M., Grow, A., Maguire, L., Wadkins, J., Richey, J., Waterston, R., and Wilson, R.	
TITLE	USDA-WashU Neospora EST Project	
JOURNAL	Unpublished (2000)	
COMMENT	Contact: Sandy Clifton, Ph.D. - Neospora USDA-WashU Neospora EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: estowatson.wustl.edu Contact David Sibley (toxoeat@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability. Seq primer: -40RP from Gibco.	
FEATURES	Location/Qualifiers	
source	1..556	
	/organism="Neospora caninum"	
	/mol_type="mRNA"	
	/db_xref="taxon:29176"	
	/dev_stage="Tachyzoite"	
	/lab_host="GC10 Competent Cells (PGC)"	
	/clone_lib="Nc-1 Tachyzoite cDNA Library 2"	
	/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed by Kelian Tang, and Robert Cole at Washington University. cDNA was synthesized from poly mRNA using an oligo-dT primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA, and products were size-selected on sephacryl S500. The cDNA were directionally cloned into the EcoRI/XhoI prepared pBluescript II SK+ vector, and electroporated into GC10 Competent Cells (PGC). The library may contain a small percentage of host or bacterial contaminants."	
ORIGIN		
Query Match	82.9%; Score 17.4; DB 7; Length 585;	
Best Local Similarity	94.7%; Pred. No. 8e+02;	
Matches	18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	3 GAGATAGGCACACAAATGA 21	
DB	419 GAGATAGGCACACAAATGA 401	
RESULT 25		
CF423038/c		
LOCUS	600 bp mRNA linear EST 03-SEP-2003	
DEFINITION	NCST3d65a09.y1 Nc-LIV Tachyzoite cDNA Library Neospora caninum	
ACCESSION	CDNA 5', mRNA sequence.	
VERSION	CF423038	
KEYWORDS	EST.	
SOURCE	CF423038.1 GI:34435739	
ORGANISM	Neospora caninum	
REFERENCE	1 (bases 1 to 600)	
AUTHORS	Cole, R., Fogarty, S., Tang, K., Howe, D. K., Sibley, L. D., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T., Theising, B.,	

Bowers, Y., Gibbons, M., Ritter, E., Bennet, J., Ronko, I., Teagareishvili, R., Fedele, M., Belaygorod, L., Franklin, C., Carr, L. M., Grow, A., Maguire, L., Wadkins, J., Richey, J., Waterston, R. and Wilson, R.

USDA-WashU Neospora EST Project
Unpublished (2000)

Contact: Sandy Clifton, Ph.D. - Neospora
USDA-WashU Neospora EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.edu

Contact David Sibley (dsibley@borcim.wustl.edu) for further information relating to organism, libraries, or clone availability.
Seq primer: -40RP from Gibco.

FEATURES

source

```
1. .600
/organism="Neospora caninum"
/mol_type="mRNA"
/db_xref="taxon:29176"
/dev_stage="Tachyzoite"
/lab_host="Electroten Blue cells (Stratagene)"
/clone_lib="Nc-Liv Tachyzoite cDNA Library"
/notes="Vector: pBluescript II SK+; Vector type: plasmid; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed by Keliang Tang, and Robert Cole at Washington University. cDNA was synthesized from Poly(A)+ mRNA using an oligo-d(T) primer containing a XhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA, and products were size-selected on sephacryl S500. The cDNA were directionally cloned into the EcoRI/XhoI prepared pBluescript II SK+ vector, and electroporated into Electroten Blue cells (Stratagene). The library may contain a small percentage of host or bacterial contaminants."
```

ORIGIN

```
Query Match      82.9%; Score 17.4; DB 7; Length 600;
Best Local Similarity 94.7%; Pred. No. 8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 3 GAGATAGGCACACCAATGA 21

```
|||||
Db 319 GAGATAGGCACACCAATGA 301
```

RESULT 26

CF416715

LOCUS

```
CF416715 602 bp mRNA linear EST 23-SEP-2003
Lr LC1ED 15R06 SAC Lumbricus rubellus Late Cocoon Library 1
Lumbricus rubellus cDNA clone Lr LC1ED 15R06 5' similar to
ref|NP 610033.1| CG10954-PA - Drosophila melanogaster, mRNA
sequence.
```

ACCESSION

CF416715.2 GI:349959873

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Lumbricus rubellus (humus earthworm)

Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;

Lumbricina; Lumbricidae; Lumbricus.

1 (bases 1 to 602)

Jones, M., Chaseley, J., Hedley, B. A., Morgan, J. C., Sturzenbaum, S., Kille, P. and Blaxter, M.

The Lumbricus rubellus EST program - Sequences from a cocoon library

Unpublished (2003)

On Sep 2, 2003 this sequence version replaced gi:34417950.

Contact: Blaxter ML

Institute of Cell, Animal and Population Biology

University of Edinburgh

Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JF, UK.

Tel: +44 131 650 6760

Fax: +44 131 670 5450

Email: mark.blaxter@ed.ac.uk

The library was prepared using protocols given by the supplier (Stratagene).

PCR Primers

FORWARD: ML3R (AGCGGATAACAATTTCACACAGGA)

BACKWARD: 17PL (CTCACTATAGGGCAATGG)

Plate: 15 row: E column: 06

Seq primer: SAC (GGGAACAAGCTGGAG)

High quality sequence stop: 504.

FEATURES

source

```
Location/Qualifiers
1. .602
/organism="Lumbricus rubellus"
/mol_type="mRNA"
/db_xref="taxon:35632"
/clone="Lr LC1ED 15R06"
/tissue_type="Whole worm"
/dev_stage="Late Cocoon"
/clone_lib="Lumbricus rubellus Late Cocoon Library 1"
/notes="Vector: pBluescript II SK+; The library was prepared using protocols given by the supplier (Stratagene)."
```

ORIGIN

```
Query Match      82.9%; Score 17.4; DB 7; Length 602;
Best Local Similarity 94.7%; Pred. No. 8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 2 TGAGATAGGCACCAACATG 20

```
|||||
Db 529 TGAGATAGGCACCAACATG 547
```

RESULT 27

BU093879

LOCUS

```
BU093879 717 bp mRNA linear EST 30-SEP-2003
Lr CA1Z 09E07 Earthworm Lambda Zap Express Library Lumbricus
rubellus cDNA clone Lr CA1Z 09E07 5' similar to pdb|IK9K|D Chain D,
Crystal Structure Of Arp23 COMPLEX, mRNA sequence.
```

ACCESSION BU093879

VERSION BU093879.1 GI:22543441

KEYWORDS EST.

SOURCE

ORGANISM

Lumbricus rubellus (humus earthworm)

Lumbricus rubellus

Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;

Lumbricina; Lumbricidae; Lumbricus.

1 (bases 1 to 717)

Sturzenbaum, S., Parkinson, J., Blaxter, M., Morgan, J., Kille, P., Schaffner, W. and Georgiev, O.

Expressed Sequence Tags from the humus earthworm L. rubellus

Unpublished (2000)

Contact: Sturzenbaum SR

School of Biosciences

Cardiff University

PO Box 911, Biomedical Building, Museum Avenue, Cardiff, CF10 3US, UK

Tel: +44 2920 874119

Fax: +44 2920 874094

Email: SturzenbaumSR@cardiff.ac.uk

The library was prepared using protocol given by supplier (Stratagene).

Plate: 09 row: E column: 07

Seq primer: T3

High quality sequence stop: 381.

Location/Qualifiers

1. .717

/organism="Lumbricus rubellus"

/mol_type="mRNA"

/db_xref="taxon:35632"

/clone="Lr CA1Z 09E07"

/tissue_type="Whole worm"

/dev_stage="adult"

/clone lib="Earthworm Lambda Zap Express Library"
/notes="Vector: pBK-CMV; The library was prepared using
protocol given by supplier (Stratagene)."

ORIGIN
Query Match 82.9%; Score 17.4; DB 5; Length 717;
Best Local Similarity 94.7%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 TGAGATAGGCACACAAATG 20

Db 598 TGAGATAGGCACACAACTAG 616

RESULT 28

BZ119502

LOCUS

DEFINITION CH230-411118.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone

CH230-411118, genomic survey sequence.

ACCESSION BZ119502

VERSION BZ119502.1 GI:23760449

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Rattus norvegicus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 904)

Zhao, S., Shetty, J., Shatman, S., Teegave, G., Geer, K.,

Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D.,

Riggs, F., de Jong, P. and Fraser, C.M.

Rat BAC End Sequences from Library CHORI-230 MboI segment

Unpublished (1999)

Other GSSs: CH230-411118.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(<http://www.chori.org/bacpac/rat230.htm>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or_ering_information.htm). BAC end

page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html

Plate: 411 row: 1 column: 18

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..904

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/strain="BN/SsNHsd/MCW"

/db_xref="taxon:10116"

/clone="CH230-411118"

/sex="Female"

/cell_type="Brain"

/clone_lib="CHORI-230 Segment 2"

/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;

CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by

Pieter de Jong"

ORIGIN

Query Match

Best Local Similarity 82.9%; Score 17.4; DB 8; Length 904;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GAGATAGGCACACAAATGA 21

Db 168 GAGATAGGCACACAAATGA 186

RESULT 29

CNS04ESL/c

LOCUS

DEFINITION

Tetraodon nigroviridis genome survey sequence

105009 of library G from Tetraodon nigroviridis, genomic survey

sequence.

AL287454

AL287454.1 GI:8025948

GSS: genome survey sequence.

Tetraodon nigroviridis

Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;

Tetraodontoidea; Tetraodontidae; Tetraodon.

1

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,

Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, P.,

Saurin, W., Bernot, A. and Weissenbach, J.

Estimate of human gene number provided by genome-wide analysis

using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)

20296633

10835645

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

CNS04ESL 919 bp DNA linear GSS 01-SEP-2000
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
105009 of library G from Tetraodon nigroviridis, genomic survey
sequence.

AL287454
AL287454.1 GI:8025948
GSS: genome survey sequence.

Tetraodon nigroviridis
Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontoidea; Tetraodontidae; Tetraodon.

1
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, P.,

Saurin, W., Bernot, A. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence

Nat. Genet. 25 (2), 235-238 (2000)
20296633
10835645

2
Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, P.,

Saurin, W., Bernot, A. and Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis

Genome Res. 10 (7), 939-949 (2000)
20359837
10899143

3 (bases 1 to 919)
Genoscope.
Direct Submission

Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr
This sequence is a single read and was generated as part of a large
scale clone-end and sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

Location/Qualifiers
1..919
/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone="105009"

/clone_lib="G"

/note="Genoscope sequence ID : C0BG105AH05SP1-end ;

PUC-ori"

ORIGIN

Query Match 82.9%; Score 17.4; DB 9; Length 919;

Best Local Similarity 85.7%; Pred. No. 8.3e+02;

Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACACAAATGA 21

Db 402 GTGCRACAGGCACACAAATGA 382

RESULT 30

BZ976367

LOCUS

DEFINITION

BZ976367 946 bp DNA linear GSS 25-MAR-2003

PUDES91TD ZM 0.6 1.0 KB Zea mays genomic clone ZMMBTa170P13,

genomic survey sequence.

ACCESSION BZ976367

VERSION BZ976367.1 GI:29203066

KEYWORDS

SOURCE

Zea mays

ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

1 (bases 1 to 946)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Bennetzen, J., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J., and

TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: T_F

Class: sheared ends.

Location/Qualifiers

1..946

/organism="Zea mays"

/mol_type="genomic DNA"

/strain="B73"

/db_xref="taxon:4577"

/clone="ZMWBtal170P13"

/clone_lib="Zm 0.6-1.0 kb"

/notes="vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high

COT selected genomic DNA library"

ORIGIN

Query Match 82.98; Score 17.4; DB 8; Length 946;
Best Local Similarity 94.78; Pred. No. 8.4e+03;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTGAGTAGGCACACAAAT 19

Db 749 GTGAGTAGGCACACAAAT 767

RESULT 31

CK173441

LOCUS CK173441 743 bp mRNA linear EST 01-JUL-2004

DEFINITION EST762761 BEA Boophilus microplus cDNA clone BEAA592, mRNA

sequence.

CK173441

VERSION CK173441.1 GI:49553980

KEYWORDS EST.

SOURCE Boophilus microplus (southern cattle tick)

ORGANISM Boophilus microplus

Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Boophilus.

1 (bases 1 to 743)

Nene, V., Quackenbush, J., George, J., and Guerrero, P.

An index of genes transcribed in the tick Boophilus microplus

Unpublished (2004)

Other ESTs: EST762762

Contact: Vishvanath Nene

Parasite Genomics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-610-5968

Fax: 301-838-0208

Email: nene@tigr.org

Seq primer: M13 forward.

Location/Qualifiers

1..743

/organism="Boophilus microplus"

/mol_type="mRNA"

/db_xref="taxon:6941"

/clone="BEAA592"

/dev_stage="Adult and larvae"

/lab_host="E. coli strain DH10B-Tona"

/clone_lib="BEA"

/note="Organ: Whole ticks and dissected organs; Vector: pExpress 1; A normalized cDNA library was custom built by Express Genomics. Oligo-dT primed cDNA was directionally cloned into NotI-EcoRV site of pExpress 1. Universal M13 primers were used to generate 5' and 3' EST data."

ORIGIN

Query Match 81.0%; Score 17; DB 7; Length 743;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGATAGGCACACAAATG 20

Db 66 AGATAGGCACACAAATG 82

RESULT 32

CK173442/c

LOCUS CK173442 743 bp mRNA linear EST 01-JUL-2004

DEFINITION EST762762 BEA Boophilus microplus cDNA clone BEAA592, mRNA

sequence.

CK173442

VERSION CK173442.1 GI:49553981

KEYWORDS EST.

SOURCE Boophilus microplus (southern cattle tick)

ORGANISM Boophilus microplus

Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;

Parasitiformes; Ixodida; Ixodoidea; Ixodidae; Boophilus.

1 (bases 1 to 743)

Nene, V., Quackenbush, J., George, J., and Guerrero, P.

An index of genes transcribed in the tick Boophilus microplus

Unpublished (2004)

Other ESTs: EST762761

Contact: Vishvanath Nene

Parasite Genomics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-610-5968

Fax: 301-838-0208

Email: nene@tigr.org

Seq primer: M13 reverse.

Location/Qualifiers

1..743

/organism="Boophilus microplus"

/mol_type="mRNA"

/db_xref="taxon:6941"

/clone="BEAA592"

/dev_stage="Adult and larvae"

/lab_host="E. coli strain DH10B-Tona"

/clone_lib="BEA"

/note="Organ: Whole ticks and dissected organs; Vector: pExpress 1; A normalized cDNA library was custom built by

Express Genomics. Oligo-dT primed cDNA was directionally

cloned into NotI-EcoRV site of pExpress 1. Universal M13

primers were used to generate 5' and 3' EST data."

ORIGIN

Query Match 81.0%; Score 17; DB 7; Length 743;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AGATAGGCACACAAATG 20

Db 678 AGATAGGCACACAAATG 662

RESULT 33

CG951752

LOCUS CG951752 895 bp DNA linear GSS 15-DEC-2003

DEFINITION MBEJC69TF mth2 Medicago truncatula genomic clone 66L17, genomic

survey sequence.

CG951752

ACCESSION CG951752

VERSION CG951752.1 GI:39863562

```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

GSS.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 895)
Town, C.D., Shetty, J., Koo, H. and Feldblyum, T.F.
Sequencing of BAC ends from Medicago truncatula
Unpublished (2003)
Other_GSSs: MBEJCG9TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
Seq primer: TGTAAACGACGCCAGT
Class: BAC ends.
FEATURES
source
1..895
Location/Qualifiers
/organism="Medicago truncatula"
/mol_type="genomic DNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="66L17"
/clone_lib="mth2"
/notes="Vector: pBeloBAC11; Site 1: HindIII; Site 2:
HindIII; Cook, D.R. and Kim, D.J., unpublished"

ORIGIN
Query Match 81.0%; Score 17; DB 9; Length 895;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 GATAGGCACCAATGA 21
|||||
Db 633 GATAGGCACCAATGA 649

RESULT 34
AZ058773/c
LOCUS
DEFINITION
158 bp DNA linear GSS 30-MAR-2000
RPCI-23-427K4.TJ RPCI-23 Mus musculus genomic clone RPCI-23-427K4,
genomic survey sequence.
ACCESSION
AZ058773
VERSION
AZ058773.1 GI:7350010
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 158)
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levine, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-427K4.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html

GSS.
Medicago truncatula (barrel medic)
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 895)
Town, C.D., Shetty, J., Koo, H. and Feldblyum, T.F.
Sequencing of BAC ends from Medicago truncatula
Unpublished (2003)
Other_GSSs: MBEJCG9TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
Seq primer: TGTAAACGACGCCAGT
Class: BAC ends.
FEATURES
source
1..158
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-427K4"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/notes="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney_and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methyase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN
Query Match 80.0%; Score 16.8; DB 8; Length 158;
Best Local Similarity 90.0%; Pred. No. 1.4e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGAGATAGGCACCAATG 20
|||||
Db 138 GTGATAGGCATACATG 119

RESULT 35
BY006779
LOCUS
DEFINITION
415 bp mRNA linear EST 06-DEC-2002
BY006779 RIKEN full-length enriched, adult male kidney Mus musculus
cDNA clone F530208M02 5', mRNA sequence.
ACCESSION
BY006779
VERSION
BY006779.1 GI:26067028
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 415)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Kikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gibsi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Malcais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shinada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Vetardo, R., Wagner, L., Wahlstedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Harai, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

```


JOURNAL Nature 420, 563-573 (2002)
 MEDLINE 22354683
 PUBMED 12466851
 COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Iotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
 Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in Riken
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

LOCUS HS_2210_B1_B11_MR CIT Approved Human Genomic Sperm Library D Homo
 sapiens genomic clone Plate=2210 Col=21 Row=D, genomic survey
 sequence.
 ACCESSION AQ338757
 VERSION AQ338757
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 456)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 1049764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones may be purchased from Research Genetics (info@resgen.com).
 BAC end Web Server: http://www.htsc.washington.edu
 Plate: 2210 row: D column: 21
 Seq primer: M13 Reverse
 Class: BAC ends
 High quality sequence stop: 456.
 Location/Qualifiers
 1. .456
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=2210 Col=21 Row=D"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
 E-Coli DH10B"

FEATURES
 source
 1. .456
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="Plate=2210 Col=21 Row=D"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
 E-Coli DH10B"

ORIGIN
 Query Match 80.0%; Score 16.8; DB 8; Length 456;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TGAGATAGGCACACAAATGA 21
 ||||| ||||| ||||| |||||
 Db 45 TGAGATAGGCACACAAATGA 26

RESULT 37
 AQ819355/c
 LOCUS HS_5295_A2_G07 SP6E RPCI-11 Human Male BAC library Homo sapiens
 genomic clone Plate=871 Col=14 Row=M, genomic survey sequence.
 DEFINITION
 ACCESSION AQ819355
 VERSION AQ819355.1 GI:5781671
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 458)
 AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

Query Match 80.0%; Score 16.8; DB 5; Length 415;
 Best Local Similarity 90.0%; Pred. No. 1.6e+03;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 TGAGATAGGCACACAAATGA 21
 ||||| ||||| ||||| |||||
 Db 387 TGAGACAGGCACACACAGTGA 406

RESULT 36
 AQ338757/c

This Page Blank (uspto)

```

Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
Db 16 ATCGTCAG 9

RESULT 12
US-09-371-772B-7010/c
; Sequence 7010, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyne Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; TITLE OF INVENTION: Levels of Vascular Endothelial Growth Factor Receptor
; FILE REFERENCE: MBH00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7010
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-371-772B-7010

Query Match 100.0%; Score 8; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCGTCAG 8
Db 13 ATCGTCAG 6

RESULT 13
US-08-584-040-3932/c
; Sequence 3932, Application US/08584040
; Patent No. 6346398
; GENERAL INFORMATION:
; APPLICANT: Pavco, Pamela
; APPLICANT: McSwiggen, James
; APPLICANT: Stinchcomb, Dan T.
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: METHOD AND REAGENT FOR THE
; TITLE OF INVENTION: TREATMENT OF DISEASES OR
; TITLE OF INVENTION: CONDITIONS RELATED TO LEVELS
; TITLE OF INVENTION: OF VASCULAR ENDOTHELIAL
; TITLE OF INVENTION: GROWTH FACTOR
; NUMBER OF SEQUENCES: 8502
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1

```

; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 3933:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-584-040-3933

Query Match 100.0%; Score 8; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
| | | | | | | |
Db 8 ATCGTCAG 1

RESULT 15

US-09-371-772B-1699/c
; Sequence 1699, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MHB00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1996-01-08
; PRIOR APPLICATION NUMBER: US 08/584,040
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1699
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-371-772B-1699

Query Match 100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
| | | | | | | |
Db 17 ATCGTCAG 10

RESULT 16

US-09-371-772B-1700/c
; Sequence 1700, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MHB00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26

; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1700
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-371-772B-1700

Query Match 100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
| | | | | | | |
Db 8 ATCGTCAG 1

RESULT 17

US-09-371-772B-6326/c
; Sequence 6326, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MHB00,876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6326
; LENGTH: 17
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-371-772B-6326

Query Match 100.0%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
| | | | | | | |
Db 15 ATCGTCAG 8

RESULT 18

US-09-866-108A-8750
; Sequence 8750, Application US/09866108A
; Patent No. 6566188
; GENERAL INFORMATION:
; APPLICANT: GU, Yizhong
; APPLICANT: JI, Yonggang
; APPLICANT: PENN, Sharon G.
; APPLICANT: HANZEL, David K.
; APPLICANT: RANK, David R.
; APPLICANT: CHEN, Wensheng
; APPLICANT: SHANNON, Mark
; TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
; FILE REFERENCE: AOMICA-7
; CURRENT APPLICATION NUMBER: US/09/866,108A
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26

```
; APPLICANT: SEGURO, KATSUYA
; TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE
; FILE REFERENCE: 0010-0937-0
; CURRENT APPLICATION NUMBER: US/09/448,310
; CURRENT FILING DATE: 1999-11-24
; PRIOR APPLICATION NUMBER: 09/109,063
; PRIOR FILING DATE: 1998-07-02
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CODON FOR
; OTHER INFORMATION: N-TERMINAL FRAGMENT
US-09-448-310-61

Query Match 100.0%; Score 8; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 9 ATCGTCAG 2

RESULT 9
US-09-884-948-61/c
; Sequence 61, Application US/09884948
; Patent No. 6821763
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, KEIICHI
; APPLICANT: NAKAMURA, NAMI
; APPLICANT: MIWA, TETSUYA
; APPLICANT: SEGURO, KATSUYA
; TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE
; FILE REFERENCE: 0010-0937-0
; CURRENT APPLICATION NUMBER: US/09/884,948
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: 09/448,310
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:CODON FOR
; OTHER INFORMATION: N-TERMINAL FRAGMENT
US-09-884-948-61

Query Match 100.0%; Score 8; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 9 ATCGTCAG 2

RESULT 10
US-08-569-150A-14
; Sequence 14, Application US/08569150A
; Patent No. 5925804
; GENERAL INFORMATION:
; APPLICANT: Hoekema, Andreas
; APPLICANT: Pen, Jan
; APPLICANT: Does, Mirjam P
; APPLICANT: Van Den Elzen, Petrus J. M
; TITLE OF INVENTION: PRODUCTION OF TREHALOSE IN PLANTS
; NUMBER OF SEQUENCES: 21
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ladass & Parry
; STREET: 26 West 61st Street
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/4" disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: WordPerfect for Windows
; SOFTWARE: WordPerfect 8
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,150A
; FILING DATE: 21-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP/94/02167
; FILING DATE: 30-JUNE-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Mass, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-010552-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 708-1890
; TELEFAX: (212) 246-8959
; TELEX: No. 5925804e
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: YES
US-08-569-150A-14

Query Match 100.0%; Score 8; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 2 ATCGTCAG 9

RESULT 11
US-09-371-772B-7009/c
; Sequence 7009, Application US/09371772B
; Patent No. 6566127
; GENERAL INFORMATION:
; APPLICANT: Ribozyme Pharmaceuticals, Inc.
; APPLICANT: Pavco, Pam
; APPLICANT: McSwiggen, Jim
; APPLICANT: Stinchcomb, Dan
; APPLICANT: Escobedo, Jaime
; TITLE OF INVENTION: Method and Reagent for the Treatment of Diseases or Conditions Re
; FILE REFERENCE: MEH800.876-J (237/198)
; CURRENT APPLICATION NUMBER: US/09/371,772B
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: US 60/005,974
; PRIOR FILING DATE: 1995-10-26
; PRIOR APPLICATION NUMBER: US 08/584,040
; PRIOR FILING DATE: 1996-01-08
; NUMBER OF SEQ ID NOS: 14225
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7009
; LENGTH: 16
; TYPE: RNA
; ORGANISM: Homo sapiens
US-09-371-772B-7009

Query Match 100.0%; Score 8; DB 4; Length 16;
```

```
; TITLE OF INVENTION: DNA ISOLATION METHOD
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,877A
; FILING DATE: 22-SEP-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,684
; FILING DATE: 05-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Bern D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 080394/0108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "probe for Candida albicans"
; US-08-934-877A-5

Query Match 100.0%; Score 8; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 11 ATCGTCAG 4

RESULT 6
US-08-871-678C-5/C
; Sequence 5, Application US/08871678C
; Patent No. 6180339
; GENERAL INFORMATION:
; APPLICANT: Sandhu, Gurpreet S, Kline, Bruce C
; TITLE OF INVENTION: Nucleic Acid Probes for the Detection and Identification
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Diagnostics Corporation
; STREET: 63 No. 6180339th Street
; CITY: Medfield
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,678C
; FILING DATE: 06-JUNE-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/373,127
; FILING DATE: 13-JANUARY-1995

; TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE
; FILE REFERENCE: 0010-0937-0
; CURRENT APPLICATION NUMBER: US/09/109,063
; CURRENT FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: JP 180010/1997
; EARLIER FILING DATE: 1997-07-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CODON FOR
; OTHER INFORMATION: N-TERMINAL FRAGMENT
; US-09-109-063-61

Query Match 100.0%; Score 8; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 9 ATCGTCAG 2

RESULT 7
US-09-109-063-61/c
; Sequence 61, Application US/09109063
; Patent No. 6013498
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, KEIICHI
; APPLICANT: NAKAMURA, NAMI
; APPLICANT: MIWA, TETSUYA
; APPLICANT: SEGURO, KATSUYA
; TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE
; FILE REFERENCE: 0010-0937-0
; CURRENT APPLICATION NUMBER: US/09/109,063
; CURRENT FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: JP 180010/1997
; EARLIER FILING DATE: 1997-07-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61
; LENGTH: 15
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CODON FOR
; OTHER INFORMATION: N-TERMINAL FRAGMENT
; US-09-109-063-61

Query Match 100.0%; Score 8; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 9 ATCGTCAG 2

RESULT 8
US-09-448-310-61/c
; Sequence 61, Application US/09448310
; Patent No. 6538122
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, KEIICHI
; APPLICANT: NAKAMURA, NAMI
; APPLICANT: MIWA, TETSUYA
```

```
; TITLE OF INVENTION: DNA ISOLATION METHOD
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,877A
; FILING DATE: 22-SEP-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,684
; FILING DATE: 05-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Bern D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 080394/0108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "probe for Candida albicans"
; US-08-934-877A-5

Query Match 100.0%; Score 8; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATCGTCAG 8
Db 11 ATCGTCAG 4

RESULT 6
US-08-871-678C-5/C
; Sequence 5, Application US/08871678C
; Patent No. 6180339
; GENERAL INFORMATION:
; APPLICANT: Sandhu, Gurpreet S, Kline, Bruce C
; TITLE OF INVENTION: Nucleic Acid Probes for the Detection and Identification
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Diagnostics Corporation
; STREET: 63 No. 6180339th Street
; CITY: Medfield
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS 6.2
; SOFTWARE: Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,678C
; FILING DATE: 06-JUNE-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/373,127
; FILING DATE: 13-JANUARY-1995
```


ADDRESSEE: PILLSBURY WINTHROP LLP
STREET: 1600 TYSONS BOULEVARD
CITY: McLean
STATE: VA
COUNTRY: USA
ZIP: 22102
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/308,759A
FILING DATE: 19-Jul-1999
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-308-759A-1
Query Match 100.0%; Score 8; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGTCAG 8
Db 1 ATCGTCAG 8
RESULT 3
US-08-435-684A-5/c
Sequence 5, Application US/08435684A
Patent No. 5707802
GENERAL INFORMATION:
APPLICANT: Sandhu, Gurpreet S.
APPLICANT: Kline, Bruce C.
TITLE OF INVENTION: Nucleic Acid Probes for the Detection
and Identification of Fungi
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba Corning Diagnostics Corp.
STREET: 63 No. 5707802th Street
CITY: Medfield
STATE: Massachusetts
COUNTRY: USA
ZIP: 02052
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS 6.2
SOFTWARE: Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/435,684A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:
NAME: Morgenstern, Arthur S.
REGISTRATION NUMBER: 28,244
REFERENCE/DOCKET NUMBER: CCD-180
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508 359-3836
TELEFAX: 508 359-3885
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 14
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Nucleic acid probe for Candida albicans

HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-435-684A-5
Query Match 100.0%; Score 8; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGTCAG 8
Db 11 ATCGTCAG 4
RESULT 4
US-08-373-127B-5/c
Sequence 5, Application US/08373127B
Patent No. 5763169
GENERAL INFORMATION:
APPLICANT: Sandhu, Gurpreet S.
APPLICANT: Kline, Bruce C.
TITLE OF INVENTION: Nucleic Acid Probes for the Detection
and Identification of Fungi
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba Corning Diagnostics Corp.
STREET: 63 No. 5763169th Street
CITY: Medfield
STATE: Massachusetts
COUNTRY: USA
ZIP: 02052
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette 3.5 inch, 1.44 Mb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS 6.2
SOFTWARE: Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/373,127B
FILING DATE: January 13, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Morgenstern, Arthur S.
REGISTRATION NUMBER: 28,244
REFERENCE/DOCKET NUMBER: CCD-180
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508 359-3836
TELEFAX: 508 359-3885
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 14
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Nucleic acid probe for Candida albicans
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-373-127B-5
Query Match 100.0%; Score 8; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCGTCAG 8
Db 11 ATCGTCAG 4
RESULT 5
US-08-934-877A-5/c
Sequence 5, Application US/08934877A
Patent No. 5958693
GENERAL INFORMATION:
APPLICANT: Sandhu, Gurpreet S.
APPLICANT: KLINE, Bruce C.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 11, 2005, 03:27:26 ; Search time 3.24051 Seconds
(without alignments)
4039.558 Million cell updates/sec

Title: US-09-674-277-1_COPY_400_407

Perfect score: 8
Sequence: 1 atcgtcag 8

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	8	100.0	10	4	US-09-508-753B-331
C 2	8	100.0	12	3	US-09-308-759A-1
C 3	8	100.0	14	1	US-08-435-684A-5
C 4	8	100.0	14	1	US-08-373-127B-5
C 5	8	100.0	14	2	US-08-934-877A-5
C 6	8	100.0	14	3	US-08-871-678C-5
C 7	8	100.0	15	3	US-09-109-063-61
C 8	8	100.0	15	4	US-09-448-310-61
C 9	8	100.0	15	4	US-09-884-948-61
C 10	8	100.0	16	2	US-08-569-150A-14
C 11	8	100.0	16	4	US-09-371-772B-7009
C 12	8	100.0	16	4	US-09-371-772B-7010
C 13	8	100.0	17	3	US-08-584-040-3932
C 14	8	100.0	17	3	US-08-584-040-3933
C 15	8	100.0	17	4	US-09-371-772B-1699
C 16	8	100.0	17	4	US-09-371-772B-1700
C 17	8	100.0	17	4	US-09-371-772B-6326
C 18	8	100.0	17	4	US-09-866-108A-8750
C 19	8	100.0	17	4	US-09-866-108A-8751
C 20	8	100.0	17	4	US-09-866-108A-8752
C 21	8	100.0	17	4	US-09-866-108A-8753
C 22	8	100.0	17	4	US-09-866-108A-8754
C 23	8	100.0	17	4	US-09-866-108A-8755
C 24	8	100.0	17	4	US-09-866-108A-8756
C 25	8	100.0	17	4	US-09-866-108A-8757
C 26	8	100.0	17	4	US-09-866-108A-8758
C 27	8	100.0	17	4	US-09-866-108A-8759

c 28	8	100.0	17	4	US-09-685-664B-1699	Sequence 1699, Ap
c 29	8	100.0	17	4	US-09-685-664B-1700	Sequence 1700, Ap
c 30	8	100.0	18	1	US-08-623-891-80	Sequence 80, Appl
c 31	8	100.0	18	2	US-08-460-751-20	Sequence 20, Appl
c 32	8	100.0	18	2	US-08-910-484-6	Sequence 6, Appl
c 33	8	100.0	18	3	US-08-584-040-4465	Sequence 4465, Ap
c 34	8	100.0	18	3	US-09-340-861-80	Sequence 80, Appl
c 35	8	100.0	18	3	US-09-634-262-80	Sequence 80, Appl
c 36	8	100.0	18	4	US-09-371-772B-2178	Sequence 2178, Ap
c 37	8	100.0	18	4	US-09-685-664B-2178	Sequence 2178, Ap
c 38	8	100.0	18	4	US-09-887-880A-81	Sequence 81, Appl
c 39	8	100.0	18	4	US-09-887-880A-103	Sequence 103, App
c 40	8	100.0	18	4	US-09-887-880A-145	Sequence 145, App
c 41	8	100.0	19	2	US-08-529-878B-1	Sequence 1, Appl
c 42	8	100.0	19	2	US-08-529-878B-5	Sequence 5, Appl
c 43	8	100.0	19	2	US-08-529-878B-46	Sequence 46, Appl
c 44	8	100.0	20	1	US-08-575-052-21	Sequence 21, Appl
c 45	8	100.0	20	1	US-08-614-516A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-508-753B-331/c
; Sequence 331, Application US/09508753B
; Patent No. 6544736
; GENERAL INFORMATION:
; APPLICANT: AKIRA SHIMAMOTO
; APPLICANT: YASUHIRO FURUICHI
; APPLICANT: YUKO SHIBATA
; APPLICANT: HIROKO FUNAKI
; APPLICANT: EIJI OHARA
; APPLICANT: MASAORI WATAHIKI
; TITLE OF INVENTION: Method for Synthesizing cDNA from mRNA sample
; FILE REFERENCE: 00162/HG
; CURRENT APPLICATION NUMBER: US/09/508,753B
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: JP 9/270324
; PRIOR FILING DATE: 1997-09-18
; NUMBER OF SEQ ID NOS: 472
; SEQ ID NO 331
; LENGTH: 10
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-508-753B-331

Query Match	100.0%;	Score 8;	DB 4;	Length 10;
Best Local Similarity	100.0%;	Pred. No. 1.3e+04;		
Matches	8;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	ATCGTCAG 8		
Db	9	ATCGTCAG 2		

RESULT 2
US-09-308-759A-1
; Sequence 1, Application US/09308759A
; Patent No. 6391593
; GENERAL INFORMATION:
; APPLICANT: Weston, Anthony
; Assenberg, Rene
; Marsh, Peter
; Mock, Graham A.
; Ray, Trevor D.
; Wharam, Susan D.
; Cady, Donald L.N.
; TITLE OF INVENTION: Modified Nucleic Acid Probes and Uses Thereof
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS: